Semimartingales and Contemporary Issues in Quantitative Finance

Younes Kchia

ECOLE POLYTECHNIQUE

Ecole Doctorale: Mathématiques et Informatique

THESE

pour obtenir le titre de

Docteur de l'Ecole Polytechnique

Spécialité : mathématiques appliquées

Younes KCHIA

Semimartingales et Problématiques Récentes en Finance Quantitative

(Semimartingales and Contemporary Issues in Quantitative Finance)

Directeur de thèse: Philip PROTTER

préparée au CMAP (Ecole Polytechnique), ORIE (Cornell University) et Statistics Department (Columbia University)

Jury :

Directeur :	Philip PROTTER	-	Columbia University
Rapporteurs :	Hans FÖLLMER	-	Humboldt Universität zu Berlin
	Monique JEANBLANC	-	Université d'Evry
Examinateurs :	Nicole EL KAROUI	-	Université Paris VI
	Jean JACOD (Président)	-	Université Paris VI
	Alex LIPTON	-	Bank of America Merrill Lynch
	Nizar TOUZI	-	Ecole Polytechnique

Acknowledgments

First and for most, I would like to thank my advisor Philip Protter. It has been an honor and a great pleasure to be his PhD student. I would like to record my gratitude to Philip for his supervision, advice and guidance through every stage of my PhD. Above all, he was equally supportive and encouraging and I truly appreciated his availability whenever his time and advice were needed. His wide mathematical culture together with his intuition (which turns out to be usually right) and uncountable ideas were exceptionnally inspiring and stimulating and gave me the chance to work on various projects, both theoretical and applied. He nourished my intellectual appetite and helped make my PhD experience productive and very enjoyable. Last but not least, he provided the excellent example of a brillant mathematician, a successful professor, a thoughtful friend and a generous man.

I am very grateful to Nizar Touzi, thanks to whom this PhD experience, although unusual, was possible. Not only did he make it possible, but enjoyable. His advices were timely and precious and I would also like to thank him for his understanding and support, both moral and financial. He made it possible for me to visit Philip for more than two semesters, at Cornell University and Columbia University and supported any initiative I could take (attending the AMAMEF conference among other things).

I gratefully acknowledge Robert Jarrow for his advice and contribution that made the research on real time bubbles detection and on discretely sampled variance swaps versus their continuous approximation interesting from more than one aspect. His expertise in economic matters was crucial in putting into context many theoretical results. I am also grateful for references he suggested and for his help in obtaining the data for the dot-com bubble companies. I take this opportunity to also thank Peter Carr and Arun Verma for their help in obtaining quickly the tick data for LinkedIn.

I gratefully thank Nicole El Karoui, Hans Föllmer, Jean Jacod, Monique Jeanblanc, Alex Lipton, Philip Protter and Nizar Touzi for their time. I am thankful and honored that despite their busy schedules, they accepted to be members of my oral defense committee. I am also particularly grateful to Monique and Hans for accepting to report this thesis.

I am also grateful to Monique for her interest in my thesis during the final stage of this PhD, for her insightful comments and the interesting references she suggested. Her work (together with her co-authors) was also an inspiring source for the work on linking progressive and initial filtration expansions (in the third section of the first chapter) and for the work on information induced credit contagion under the conditional density assumption with multiple non-ordered defaults (in the third section of the second chapter).

I would like to thank my friend and colleague Martin Larsson. As a colleague, I thank him deeply and gratefully acknowledge his collaboration all along my PhD. His ideas were inspiring and his enthusiasm contagious and motivational. I had a lot of fun working together and very much enjoyed our long productive working hours, in Ithaca, Paris, New

York and Bled. I really hope we will keep up our collaboration in the future. As a friend, I thank him a lot for his hospitality during my visits to Ithaca and for his wise advices.

I convey special acknowledgement to Alex Lipton and the Commodities Quant team at Bank of America - Merrill Lynch who welcomed me three times during my Master and PhD as an intern. These experiences were helpful to keep in mind the issues practitionners face and provided me with a valuable and complementary work experience. Many thanks to David, Jean-François, John, Stewart, Ying and Yuanbo for being welcoming and for their help and their time. A special thank to David, for his help with C^{++}. I take also this opportunity to thank two other quants, Dominique and Pascal, who showed a vivid interest to my research.

My time at Cornell was made enjoyable in large part due to Sophia and Martin, to whom I am indebted more than they know. Also, I would like to thank Tia, Matt, Rolf and Colin among many other graduate students who made a winter in Ithaca very nice and the time spent at Rhodes interesting.

The semester I spent at Columbia was enriched by the graduate students I shared an office with and the interesting discussions during Philip's class with Radka, Johannes, Micheal and Murad. The New York experience could not be greater thanks to Mariem's visit and my amazing flatmates Charlotte, Ahcène, Aurélie and Alfred. Thank you all for the priceless memories.

The short amount of time I spent at the CMAP was made enjoyable thanks to my office mates. I keep a good memory of the coffee breaks with Maisem and I am thankful to Dylan and Réda for their help with my TA duties when I was abroad and for the interesting discussions we had and ideas we shared on our different research topics. I would like to take this opportunity to also thank Caroline for her previous advices.

I would like to thank Nasséra at the Ecole Polytechnique for her indispensable help dealing with funding my travels back and forth to the US. I am also grateful to Kathy at Cornell University and Dood at Columbia University for their help dealing with the administrative matters during my stays in the US.

Words fail to express all my gratitude and appreciation to my family for all their love and encouragement. To my parents, M'hamed and Ouafa, for their continued prayers and for raising me with a love of science and supporting me in all my pursuits. To Ali and Momo, for being such supportive and caring brothers. To my sisters and brothers in law, Leila, Loubna and Pierre, for their encouragement and the great time we spent together. And most of all, to my loving and patient wife, Mariem, whose dedication, support and confidence in me, has taken the load off my shoulders. Thank you.

Finally, I would like to thank everybody who helped consciously or unconsciously to the successful realization of my thesis and I express my apologies that I could not mention everyone.

Semimartingales and Contemporary Issues in Quantitative Finance

Abstract: In this thesis, we study various contemporary issues in quantitative finance.

The first chapter is dedicated to the stability of the semimartingale property under filtration expansion. We study first progressive filtration expansions with random times. We show how semimartingale decompositions in the expanded filtration can be obtained using a natural link between progressive and initial expansions. The link is, on an intuitive level, that the two coincide after the random time. We make this idea precise and use it to establish known and new results in the case of expansion with a single random time. The methods are then extended to the multiple time case, without any restrictions on the ordering of the individual times. We then look to the expanded filtrations from the point of view of filtration shrinkage. We turn then to studying progressive filtration expansions with processes. Using results from the weak convergence of σ-fields theory, we first establish a semimartingale convergence theorem, which we apply in a filtration expansion with a process setting and provide sufficient conditions for a semimartingale of the base filtration to remain a semimartingale in the expanded filtration. A first set of results is based on a Jacod's type criterion for the increments of the process we want to expand with. An application to the expansion of a Brownian filtration with a time reversed diffusion is given through a detailed study and some known examples in the litterature are recovered and generalized. Finally, we focus on filtration expansion with continuous processes and derive two new results. The first one is based on a Jacod's type criterion for the successive hitting times of some levels and the second one is based on honest times assumptions for these hitting times. We provide examples and see how those can be used as first steps toward harmful dynamic insider trading models. In the expanded filtration the finite variation term of the price process can become singular and arbitrage opportunities (in the sense of FLVR) can therefore arise in these models.

In the second chapter, we reconcile structural models and reduced form models in credit risk from the perspective of the information induced credit contagion effect. That is, given multiple firms, we are interested on the behaviour of the default intensity of one firm at the default times of the other firms. We first study this effect within different specifications of structural models and different levels of information. Since almost all examples are non tractable and computationally very involved, we then work with the simplifying assumption that conditional densities of the default times exist. The classical reduced-form and filtration expansion framework is therefore extended to the case of multiple, non-ordered defaults times having conditional densities. Intensities and pricing formulas are derived, revealing how information-driven default contagion arises in these models. We then analyze the impact of ordering the default times before expanding the filtration. While not important for pricing, the effect is significant in the context of risk management, and becomes even more pronounced for highly correlated and asymmetrically distributed defaults. We provide a general scheme for constructing and simulating the default times, given that a model for the conditional densities has been chosen. Finally, we study particular conditional density models and the information induced credit contagion effect within them.

In the third chapter, we provide a methodology for a real time detection of bubbles. After the 2007 credit crisis, financial bubbles have once again emerged as a topic of current concern. An open problem is to determine in real time whether or not a given asset's price process exhibits a bubble. Due to recent progress in the characterization of asset price bubbles using the arbitrage-free martingale pricing technology, we are able to propose a new methodology for answering this question based on the asset's price volatility. We limit ourselves to the special case of a risky asset's price being modeled by a Brownian driven stochastic differential equation. Such models are ubiquitous both in theory and in practice. Our methods use non parametric volatility estimation techniques combined with the extrapolation method of reproducing kernel Hilbert spaces. We illustrate these techniques using several stocks from the alleged internet dot-com episode of 1998 - 2001, where price bubbles were widely thought to have existed. Our results support these beliefs. During May 2011, there was speculation in the financial press concerning the existence of a price bubble in the aftermath of the recent IPO of LinkedIn. We analyzed stock price tick data from the short lifetime of this stock through May 24, 2011, and we found that LinkedIn has a price bubble.

The last chapter is about discretely sampled variance swaps, which are volatility derivatives that trade actively in OTC markets. To price these swaps, the continuously sampled approximation is often used to simplify the computations. The purpose of this chapter is to study the conditions under which this approximation is valid. Our first set of theorems characterize the conditions under which the discretely sampled variance swap values are finite, given the values of the continuous approximations exist. Surprisingly, for some otherwise reasonable price processes, the discretely sampled variance swap prices do not exist, thereby invalidating the approximation. Examples are provided. Assuming further that both variance swap values exist, we study sufficient conditions under which the discretely sampled values converge to their continuous counterparts. Because of its popularity in the literature, we apply our theorems to the 3/2 stochastic volatility model. Although we can show finiteness of all swap values, we can prove convergence of the approximation only for some parameter values.

Keywords: Initial and progressive filtration expansions, progressive filtration expansion with processes, weak convergence of σ-fields, convergence of semimartingales, dynamic models for insider trading, compensators of random times, multiple non ranked random times, conditional density assumption, credit contagion, structural models, reduced form models, martingale theory of asset bubbles, bubbles detection, strict local martingales, non parametric volatility estimation, RKHS extrapolation, continuously/discretely sampled variance swaps.

Résumé: Dans cette thèse, nous étudions différentes problématiques d'actualité en finance quantitative.

Le premier chapitre est dédié à la stabilité de la propriété de semimartingale après grossissement de la filtration de base. Nous étudions d'abord le grossissement progressif d'une filtration avec des temps aléatoires et montrons comment la décomposition de la semimartingale dans la filtration grossie est obtenue en utilisant un lien naturel entre la filtration grossie initiallement et celle grossie progressivement. Intuitivement, ce lien se résume au fait que ces deux filtrations coincident après le temps aléatoire. Nous précisons cette idée et l'utilisons pour établir des résultats connus pour certains et nouveaux pour d'autres dans le cas d'un grossissement de filtrations avec un seul temps aléatoire. Les méthodes sont alors étendues au cas de plusieurs temps aléatoires, sans aucune restriction sur l'ordre de ces temps. Nous étudions ensuite ces filtrations grossies du point de vue des rétrécissements des filtrations. Nous nous intéressons enfin au grossissement progressif de filtrations avec des processus. En utilisant des résultats de la convergence faible de tribus, nous établissons d'abord un théorème de convergence de semimartingales, que l'on appliquera dans un contexte de grossissement de filtrations avec un processus pour obtenir des conditions suffisantes pour qu'une semimartingale de la filtration de base reste une semimartingale dans la filtration grossie. Nous obtenons des premiers résultats basés sur un critère de type Jacod pour les incréments du processus utilisé pour grossir la filtration. Nous nous proposons d'appliquer ces résultats au cas d'un grossissement d'une filtration Brownienne avec une diffusion retournée en temps et nous retrouvons et généralisons quelques examples disponibles dans la littérature. Enfin, nous concentrons nos efforts sur le grossissement de filtrations avec un processus continu et obtenons deux nouveaux résultats. Le premier est fondé sur un critère de Jacod pour les temps d'atteinte successifs de certains niveaux et le second est fondé sur l'hypothèse que ces temps sont honnêtes. Nous donnons des examples et montrons comment cela peut constituer un premier pas vers des modèles dynamiques de traders initiés donnant naissance à des opportunités d'arbitrage nocives. Dans la filtration grossie, le terme à variation finie du processus de prix peut devenir singulier et des opportunités d'arbitrage (au sens de FLVR) apparaissent clairement dans ces modèles.

Dans le deuxième chapitre, nous réconcilions les modèles structuraux et les modèles à forme réduite en risque de crédit, du point de vue de la contagion de crédit induite par le niveau d'information disponible à l'investisseur. Autrement dit, étant données de multiples firmes, nous nous intéressons au comportement de l'intensité de défaut (par rapport à une filtration de base) d'une firme donnée aux temps de défaut des autres firmes. Nous étudions d'abord cet effet sous des spécifications différentes de modèles structuraux et sous différents niveaux d'information, et tirons, par l'exemple, des conclusions positives sur la présence d'une contagion de crédit. Néanmoins, comme plusieurs exemples pratiques ont un coup calculatoire élevé, nous travaillons ensuite avec l'hypothèse simplificatrice que les temps de défaut admettent une densité conditionnelle par rapport à la filtration de base. Nous étendons alors des résultats classiques de la théorie de grossissement de

filtrations avec des temps aléatoires aux temps aléatoires non-ordonnés admettant une densité conditionnelle et pouvons ainsi étendre l'approche classique de la modélisation à forme réduite du risque de crédit à ce cas général. Les intensités de défaut sont calculées et les formules de pricing établies, dévoilant comment la contagion de crédit apparaît naturellement dans ces modèles. Nous analysons ensuite l'impact d'ordonner les temps de défaut avant de grossir la filtration de base. Si cela n'a aucune importance pour le calcul des prix, l'effet est significatif dans le contexte du management de risque et devient encore plus prononcé pour les défauts très corrélés et asymétriquement distribués. Nous proposons aussi un schéma général pour la construction et la simulation des temps de défaut, étant donné qu'un modèle pour les densités conditionnelles a été choisi. Finalement, nous étudions des modèles de densités conditionnelles particuliers et la contagion de crédit induite par le niveau d'information disponible au sein de ces modèles.

Dans le troisième chapitre, nous proposons une méthodologie pour la détection en temps réel des bulles financières. Après la crise de crédit de 2007, les bulles financières ont à nouveau émergé comme un sujet d'intéret pour différents acteurs du marché et plus particulièrement pour les régulateurs. Un problème ouvert est celui de déterminer si un actif est en période de bulle. Grâce à des progrès récents dans la caractérisation des bulles d'actifs en utilisant la théorie de pricing sous probabilité risque-neutre qui caractérise les processus de prix d'actifs en bulles comme étant des martingales locales strictes, nous apportons une première réponse fondée sur la volatilité du processus de prix de l'actif. Nous nous limitons au cas particulier où l'actif risqué est modélisé par une équation différentielle stochastique gouvernée par un mouvement Brownien. Ces modèles sont omniprésents dans la littérature académique et en pratique. Nos méthodes utilisent des techniques d'estimation non paramétrique de la fonction de volatilité, combinées aux méthodes d'extrapolation issues de la théorie des *reproducing kernel Hilbert spaces*. Nous illustrons ces techniques en utilisant différents actifs de la bulle internet (dot-com bubble) de la période 1998 - 2001, où les bulles sont largement acceptées comme ayant eu lieu. Nos résultats confirment cette assertion. Durant le mois de Mai 2011, la presse financière a spéculé sur l'existence d'une bulle d'actif après l'OPA sur LinkedIn. Nous analysons les prix de cet actif en nous basant sur les données tick des prix et confirmons que LinkedIn a connu une bulle pendant cette période.

Le dernier chapitre traîte des variances swaps échantillonnés en temps discret. Ces produits financiers sont des produits dérivés de volatilité qui tradent activement dans les marchés OTC. Pour déterminer les prix de ces swaps, une approximation en temps continu est souvent utilisée pour simplifier les calculs. L'intérêt de ce chapitre est d'étudier les conditions garantissant que cette approximation soit valable. Les premiers théorèmes caractérisent les conditions sous lesquelles les valeurs des variances swaps échantillonnés en temps discret sont finies, étant donné que les valeurs de l'approximation en temps continu sont finies. De manière étonnante, les valeurs des variances swaps échantillonnés en temps discret peuvent etre infinies pour des modèles de prix raisonnables, ce qui rend la pratique de marché d'utiliser l'approximation en temps continu invalide. Des examples

sont fournis. En supposant ensuite que le payoff en temps discret et son approximation en temps continu ont des prix finis, nous proposons des conditions suffisantes pour qu'il y ait convergence de la version discrète vers la version continue. Comme le modèle à volatilité stochastique 3/2 est de plus en plus populaire, nous lui appliquons nos résultats. Bien que nous pouvons démontrer que les deux valeurs des variances swaps sont finies, nous ne pouvons démontrer la convergence de l'approximation que pour certaines valeurs des paramètres du modèle.

Mots-clés: Grossiment initial et progressif de filtration, grossissement progressif de filtrations avec des processus, convergence faible de tribus, convergence de semi-martingales, modèles dynamiques de traders initiés, compensateurs de temps aléatoires, temps aléatoires multiples non-ordonnés, hypothèse de densité conditionnelle, contagion de crédit, modèles structuraux, modèles à forme réduite, théorie martingale des bulles d'actif, détection de bulles financières, martingales locales strictes, extrapolation RKHS, variances swaps échantillonnés en temps discret ou en temps continu.

Contents

Chapter 1

Filtration expansions and semimartingales

Contents

1.1 Introduction

One of the key insights of K. Itô when he developed the Itô integral was to restrict the space of integrands to what we now call predictable processes. This allowed the integral to have a type of bounded convergence theorem that N. Wiener was unable to obtain with unrestricted random integrands. The Itô integral has since been extended to general semimartingales. If one tries however to expand (i.e. to enlarge) the filtration, then one is playing with fire, and one may lose the key properties Itô originally obtained with his restriction to predictable processes. In the 1980's a theory of such filtration expansions was nevertheless successfully developed for two types of expansion: initial expansions and progressive expansions; see for instance [72] and [92]. More recent partial expositions can be found in the following books: Yor, *Some aspects of Brownian motion, Part II* (see [130]), Dellacherie, Maisonneuve and Meyer, *Probabilités et Potentiel, chapitres XVII-XXIV* (see [37]), Protter *Stochastic integration and differential equations* (see [117, Chapter VI]) and Mansuy and Yor, *Random times and (Enlargement of) filtrations in a Brownian setting* (see [109]). These two main types of filtration expansion are well-studied topics that have been investigated both in theoretical and applied contexts. For theoretical results on initial filtration expansion, see for instance [72] and [1]. Good references for theoretical results on progressive filtration expansion are [92], [86], [116], [60], [102] among others. The subject has regained interest recently, due to applications in Mathematical Finance. We refer first the reader to [44], one of the first papers to study filtrations expansion in the setting of jump processes with applications to finance. The more recent regain of interest is examplified by the large litterature using progressive expansions of filtrations in credit risk, see for instance [87], [85], [59], [101], and by the one using initial filtration expansion in insider trading models, see for instance [2], [18], [71]. The initial expansion of a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ with a random variable τ is the filtration $\mathbb{H}$ obtained as the right-continuous modification of $(\mathcal{F}_t \vee \sigma(\tau))_{t\geq 0}$. *A priori* there is no particular interpretation attached to this random variable. The progressive expansion $\mathbb{G}$ is obtained as any right-continuous filtration containing $\mathbb{F}$ and making τ a stopping time. In this case τ should of course be nonnegative since it has the interpretation of a random time. When referring to the progressive expansion with a random variable in this thesis, we mean the smallest such filtration. One is usually interested in the cases where $\mathbb{F}$ semimartingales remain semimartingales in the expanded filtrations and in their decompositions when viewed as semimartingales in the expanded filtrations. For the initial filtration expansion $\mathbb{H} = (\mathcal{H}_t)_{t\geq 0}$ of a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ with a random variable τ, one well-known situation where this holds is when *Jacod's criterion* (which is recalled in Assumption 1) is satisfied (see [72] or alternatively [117, Theorem 10, p. 371]), and as far as one is concerned by the progressive filtration expansion $\mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$, this always holds up to the random time τ as proved by Jeulin and Yor and holds on all $[0, \infty)$ for honest times (see [92]). Apart from the previous references, we also point out the nice work by Callegaro, Jeanblanc and Zargari in [17] summarizing the results in both types of filtration expansions under a stronger assumption than Assumption 1 (equivalence of the conditional laws of τ w.r.t its law are assumed rather than just their absolute continuity). Under this same assumption, Amendinger already provided martingale representation theorems for initially enlarged filtrations in 2000 in [1].

The aim of this chapter is fourfold. In the first section, we introduce the mathematical tools that will be useful for the two first chapters of this thesis. Yor already noticed that the decompositions in the initially and progressively expanded filtrations are related and showed how one can obtain the decompositions of some $\mathbb{F}$ local martingales in the progressively expanded filtration under representability assumptions of some crucial $\mathbb{F}$ martigales related to the random time τ. We refer the reader to [129] for more details. However, these two filtrations remained then often viewed and studied in the literature one independently from the other. The purpose of the second section of the present chapter is to build further on this idea and demonstrate that these two filtration expansions are not inherently different—in fact, there is a very natural connection between the initial and progressive expansions. The reason is, on an intuitive level, that the filtrations $\mathbb{G}$ and $\mathbb{H}$ coincide after time τ. We make this idea precise for filtrations $\mathbb{H}$ that are not necessarily obtained as initial expansions. This, in combination with the classical Jeulin - Yor Theorem (see Theorem 15), allows us to show how the semimartingale decomposition of an $\mathbb{F}$ local martingale, when viewed in the progressively expanded filtration $\mathbb{G}$, can be obtained on all of $[0, \infty)$, provided that its decomposition in the filtration $\mathbb{H}$ is known. Our result is given in Theorem 16. As already mentioned, one well-known situation where this is the case is when *Jacod's criterion* is satisfied. This is, however, not the only case, and we give an example using techniques based on Malliavin calculus developed by Imkeller et al. [71]. Related issues and Malliavin calculus based techniques can also be found in [30]. These developments, which all concern expansion with a single random time, are treated in Section 1.3.1. The technique is, however, applicable in much more general situations than expansion with a single random time. As an indication of this, we perform *en passant* the same analysis for what we call the (τ, X)-progressive expansion of $\mathbb{F}$, denoted $\mathbb{G}^{(\tau,X)}$. This expansion is a particular case of a progressive filtration expansion introduced by Jeulin-Yor, where some given σ-field $\mathcal{E}$ is added at time τ. The time τ becomes a stopping time in the larger filtration, and the random variable X becomes $\mathcal{G}_\tau^{(\tau,X)}$-measurable, $\mathcal{E}$ being $\sigma(X)$. In Section 1.3.2 we extend these ideas in order to deal with the case where the base filtration $\mathbb{F}$ is expanded progressively with a whole vector $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_n)$ of random times, and we do not impose any conditions on the ordering of the individual times. After establishing a general semimartingale decomposition result (see Theorem 19 which is the primary result of this section) we treat the special case where Jacod's criterion is satisfied for the whole vector $\boldsymbol{\tau}$, and we show how the decompositions may be expressed in terms of $\mathbb{F}$ conditional densities of $\boldsymbol{\tau}$ with respect to its law. This can be compared to results in [86] where the times are assumed to be recursively initial. This type of decomposition result is also obtained for a multidimentional (τ, X)-progressive expansion of $\mathbb{F}$ in Theorem 20. Finally, in Section 1.3.3 we take a different point of view and study the link between the filtrations $\mathbb{G}$ and $\mathbb{H}$ from the perspective of filtration shrinkage. This amounts to pushing forward Yor's original idea consisting in projecting down the decomposition in $\mathbb{H}$ onto $\mathbb{G}$ to obtain the $\mathbb{G}$ decomposition. We know that the optional projection of a local martingale M onto a filtration to which it is not adapted may lose the local martingale property, see [50]. However, general conditions for when this happens are not available. In this section, theorem 21 solves partially the problem when working in a filtration expansion setting. Starting from an $\mathbb{F}$ local martingale that remains $\mathbb{H}$ semimartingale, we prove that the optional projection onto $\mathbb{G}$ of the $\mathbb{H}$ local martingale part of M is a $\mathbb{G}$ local martingale. In the case of a random time that avoids all $\mathbb{F}$ stopping times and whose $\mathbb{F}$ conditional probabilities are equivalent to some deterministic measure, Callegaro, Jeanblanc and Zargari already used the idea of projecting down the $\mathbb{H}$ decomposition of an $\mathbb{F}$ local martingale to obtain its decomposition in $\mathbb{G}$, see [17]. Some of their results follow from ours, and our technique allows to avoid the heavy manual computations involving dual predictable projections that arise naturally in this problem, see Corollary 9.

In the third section of this chapter, we go beyond the simple cases of initial expansion and progressive expansion with random variables. Instead we consider the more complicated case of expansion of a filtration through dynamic enlargement, by adding a stochastic process as it evolves simultaneously to the evolution of the original process. Jeulin already considered in [89] such enlargement and expanded the natural filtration of a Bessel 3 process with its remaining infimum. He proved that some Brownian motion of the base filtration remains a semimartingale in the expanded filtration. Ankirchner and al. studied in [5] arbitrary filtration expansions using the concept of decoupling measures. Ankirchner also considered in [4] the case of an arbitrary filtration expansion $\mathbb{G}$ of a base filtration $\mathbb{F}$ and provided sufficient conditions for a purely discontinuous $\mathbb{F}$ martingale M to be a semimartingale relative to the filtration $\mathbb{G}$. The primary aim of the third and forth sections is to provide general theorems for this kind of results to hold when expanding a filtration with a process. That is, for a given filtration $\mathbb{F}$ and a given càdlàg process X, the smallest right-continuous filtration containing $\mathbb{F}$ and to which X is adapted will be called the progressive expansion of $\mathbb{F}$ with X and we will investigate the stability of the semimartingale property of $\mathbb{F}$ semimartingales in progressive expansions of $\mathbb{F}$ with càdlàg processes X. Recall that we were able to link in section 1.3 the initial and progressive filtration expansions and to extend the usual semimartingale stability results to the multiple time case, without any restrictions on the ordering of the individual times and more importantly to the filtration expanded with a counting process $N^n_t = \sum_{i=1}^n X_i 1_{\{\tau_i \le t\}}$, i.e. the smallest right-continuous filtration containing $\mathbb{F}$ and to which the process N^n is adapted. We apply these results together with results from the theory of weak convergence of σ-fields recently developed by Antonelli, Coquet, Kohatsu-Higa, Mackevicius, Mémin, and Slominski (see for instance [7],[28],[29]) to obtain more sophisticated enlargement possibilities. We combine the convergence results with an extension of an old result of Barlow and Protter [11], finally obtaining a general criterion that guarantees that $\mathbb{F}$ semimartingales satisfying suitable integrability assumptions remain semimartingales in the expanded filtration. Our key results include the forms of the semimartingale decompositions in the enlarged filtrations. Section 1.4 extends the main theorem in [11] and proves a general result on the convergence of $\mathbb{G}^n$ special semimartingales to a $\mathbb{G}$ adapted process X, where $(\mathbb{G}^n)_{n\geq 1}$ and $\mathbb{G}$ are filtrations such that $\mathcal{G}^n_t$ converges weakly to $\mathcal{G}_t$ for each $t \geq 0$. The process X is proved to be, in Theorem 22, a $\mathbb{G}$ special semimartingale under sufficient conditions on the regularity of the local martingale and finite variation parts of the $\mathbb{G}^n$ semimartingales. This theorem is the first main result of this section. This is then applied to the case where the filtrations $\mathbb{G}^n$ are obtained by progressively expanding a base filtration $\mathbb{F}$ with processes N^n converging in probability to some process N. We provide sufficient conditions for an $\mathbb{F}$ semimartingale to remain a $\mathbb{G}$ semimartingale, where $\mathbb{G}$ is the progressive expansion of $\mathbb{F}$ with N. We apply this result to the case where the base filtration $\mathbb{F}$ is progressively expanded with a càdlàg process whose increments satisfy a *generalized Jacod's criterion* (see Assumption 2) with respect to the filtration $\mathbb{F}$ along *some* sequence of subdivisions whose mesh tends to zero. This leads to Theorem 27, the second main result of this section. An application to the expansion of a Brownian filtration with a time reversed diffusion is given through a detailed study, and the canonical decomposition of the Brownian motion in the expanded filtration is provided. A possible application to stochastic volatility models is then suggested and the example in Kohatsu-Higa [98] is easily recovered. The techniques developed in this section require a long preliminary treatment of the weak convergence of σ-fields, and to a lesser extent the weak convergence of filtrations. We treat these questions in a preliminary section (see section 1.2.4) that summarizes the mathematical tools we need. It is wise to indicate that the results of practical interest are Theorems 25 and 27, which show how one can expand filtrations with processes and have semimartingales remain semimartingales in the enlarged filtrations. We also wish to mention here that the example provided in Theorem 29 shows how the hypotheses (perhaps a bit strange at first glance) of Theorem 27 can arise naturally in applications, and it shows the potential utility of our results. That said, the preliminary results on the weak convergence of σ-fields have an interest in their own right.

Finally, section 1.5 focuses on the case where the process X used to enlarge the base filtration $\mathbb{F}$ is continuous. In this case, new criterions that are sometimes easier to check can be found to ensure that a given $\mathbb{F}$ semimartingale M, satisfying some integrability assumptions, remains a semimartingale in the expanded filtration. Now, instead of assuming that the increments of X satisfy some kind of Jacod's assumption (as in Assumption 2), we will instead work with the successive hitting times of X of some given levels. In case these times are either honest or initial (in the sense that they can be used to enlarge initially $\mathbb{F}$ and do not affect the semimartingale property of M in the expanded filtration) the same conclusions as in the previous section can be reached. This is the aim of both Theorems 36 and 37. The last result allows to recover Jeulin's example (actually this is Pitman's theorem that Jeulin proved using filtration expansions techniques), where the natural filtration of a Bessel 3 process is progressively expanded with its remaining infimum. We also extend this result to the case where we expand the natural filtration of a transient diffusion R with its remaining infimum. Many authors studied arbitrage opportunities within insider trading models (see for instance [98], [70] and [71]). In our last example, we notice that the finite variation part in the progressively expanded filtration is now singular w.r.t Lebesgue measure. This provides a dynamic insider trading model with unrisky arbitrage opportunities. We hope that these ideas can constitute first steps toward dynamic insider models, which would be closer to reality.

1.2 Mathematical preliminaries

In this section, we recall briefly some basic results from the general theory of stochastic processes. The well known theorems in initial and progressive filtration expansion are then given in few details. We emphasize two types of results: the semimartingale decompositions in the expanded filtrations of a semimartingale in the base filtration, and the compensator of a random time τ in the progressive expansion of a filtration with τ, as given by Jeulin-Yor Theorem. Up to this point, we do not provide proofs since all these results are well known and an excellent and much more exhaustive review can be found in [114] or [117]. We then provide few extensions to filtration expansions with counting processes and derive some lemmas useful for section 1.5. Finally, and as already mentioned in the introduction, we provide preliminary results on the weak convergence of σ-fields and of filtrations which is the crucial tool needed in section 1.4.

Throughout this section, we assume we are given a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ that satisfies the "usual hypotheses", i.e. $\mathbb{F}$ is right continuous and $\mathcal{F}_0$ contains all the P null sets of $\mathcal{F}$. Therefore $\mathcal{F}_t$ contains all the P null sets of $\mathcal{F}$ as well, for any $t \geq 0$. We will also assume throughout this section that all filtrations that we consider satisfy the usual hypotheses.

1.2.1 General theory of stochastic processes

In this subsection, we mainly recall the fundamental notions of optional and predictable projections and the dual optional and predictable projections. First, recall that the optional σ-field $\mathcal{O}$ is the σ-algebra defined on $\mathbb{R}^+ \times \Omega$ generated by all càdlàg (right-continuous with left limits) processes that are $\mathbb{F}$ adapted, and that the predictable σ-field $\mathcal{P}$ is the σ-algebra defined on $\mathbb{R}^+ \times \Omega$ generated by all càg (left-continuous) processes that are $\mathbb{F}$ adapted. A time τ is said to be predictable if the stochastic interval $[\![0, \tau[\![$ is predictable.

1.2.1.1 Optional and predictable projections

Theorem 1 *Let X be a measurable process either positive or bounded.*

(i) There exists a unique up to indistinguishability optional process, called the optional projection of X and denoted ${}^oX$, such that

$$E(X_T 1_{\{T<\infty\}} \mid \mathcal{F}_T) = {}^oX_T 1_{\{T<\infty\}} \quad a.s.$$

for every stopping time T.

(ii) There exists a unique up to indistinguishability predictable process, called the predictable projection of X and denoted ${}^pX$, such that

$$E(X_T 1_{\{T<\infty\}} \mid \mathcal{F}_{T-}) = {}^pX_T 1_{\{T<\infty\}} \quad a.s.$$

for every predictable stopping time T.

The following characterization for predictable processes appears to be very useful.

Theorem 2 *If X is càdlàg adapted process, then X is predictable if and only if $\Delta X_T = 0$ a.s. on $\{T < \infty\}$ for all totally inaccessible stopping times T and $X_T 1_{\{T<\infty\}}$ is $\mathcal{F}_{T-}$ measurable for every predictable stopping time T.*

1.2.1.2 Dual optional and predictable projections

If the projection operation defines rigourously what we would write formally as $E(X_t \mid \mathcal{F}_t)$, the dual projection allows to define rigourously the quantity $\int_0^t E(dA_s \mid \mathcal{F}_s)$ for a right-continuous increasing integrable process A, which is not necessary adapted.

Theorem 3 *Let A be a right-continuous increasing integrable process. Then*

(i) There exists a unique up to indistinguishability optional increasing process A^o, called dual optional projection of A, such that

$$E(\int_0^\infty X_s dA_s^o) = E(\int_0^\infty {}^oX_s dA_s)$$

for every bounded measurable process X.

(ii) There exists a unique up to indistinguishability predictable increasing process A^p, called dual predictable projection of A, such that

$$E(\int_0^\infty X_s dA_s^p) = E(\int_0^\infty {}^pX_s dA_s)$$

for every bounded measurable process X.

The jumps of the dual predictable projection are given by the following lemma.

Lemma 1 *Let T be a predictable stopping time. Then $\Delta A_T^p = E(\Delta A_T \mid \mathcal{F}_{T-})$ a.s. with the convention $\Delta A_\infty^p = 0$*

If A is supposed to be optional, then A^p is the compensator of A.

Lemma 2 *Let A be an optional process of integrable variation. Then A^p is the unique increasing predictable process such that $A - A^p$ is a martingale.*

Finally, we will also need the following result in section 1.3.3.

Theorem 4 *(i) Every predictable process of finite variation is locally integrable.*

(ii) An optional process A of finite variation is locally integrable if and only if there exists a predictable process $\tilde{A}$ of finite variation such that $A - \tilde{A}$ is a local martingale which vanishes at 0. When it exists, $\tilde{A}$ is unique. We say that $\tilde{A}$ is the predictable compensator of A.

1.2.1.3 Azéma supermartingales and dual projections associated with random times

The following processes will be of crucial importance in filtration expansions with random times. Let τ be a random time and Z the optional projection onto $\mathbb{F}$ of $1_{[\![0,\tau[\![}$. Then Z is an $\mathbb{F}$ supermartingale given by a càdlàg version of $P(\tau > t \mid \mathcal{F}_t)$ and is called Azéma supermartingale. Let a and A be respectively the dual optional and predictable projections of the process $1_{\{\tau \leq t\}}$. Then $\nu_t = E(a_\infty \mid \mathcal{F}_t) = a_t + Z_t$ is a BMO martingale. Finally let $Z_t = M_t - A_t$ be the Doob-Meyer decomposition of the supermartingale Z. The following holds.

Lemma 3 *(i) If τ avoids all $\mathbb{F}$ stopping times, then $A_t = a_t$ is continuous.*

(ii) If all $\mathbb{F}$ martingales are continuous, a is predictable and consequently $A = a$.

(iii) Under the two conditions above, Z is continuous.

The aim of the two next sections is to present the theory of filtration expansions. The question of interest is the one we already emphasized in the introduction. Do $\mathbb{F}$ semimartingales remain semimartingales in a filtration $\mathbb{G}$ containing $\mathbb{F}$?

Definition 1 *We say that hypothesis $(H^{'})$ holds between $\mathbb{F}$ and $\mathbb{G}$ if every $\mathbb{F}$ semimartingale is a $\mathbb{G}$ semimartingale.*

When hypothesis $(H^{'})$ does not hold between $\mathbb{F}$ and $\mathbb{G}$ we focus on finding conditions under which a given $\mathbb{F}$ semimartingale is a $\mathbb{G}$ semimartingale. Of course, one can restrict itself to $\mathbb{F}$ martingales. The reverse situation is known as filtration shrinkage. We do not cite the known results on filtration shrinkage here, but recall them throughout the thesis when needed. We refer to [50] for a nice paper on filtration shrinkage issues. However, we do recall Stricker's theorem which we will intensively use in the next sections.

Theorem 5 (Stricker) *Let $\mathbb{F} \subset \mathbb{G}$ two filtrations. If X is a $\mathbb{G}$ semimartingale which is $\mathbb{F}$ adapted, then X is also an $\mathbb{F}$ semimartingale.*

As already mentioned in the introduction, there are essentially two types of filtration expansions: initial and progressive filtration expansion.

1.2.2 Initial filtration expansion with a random variable

Let ξ be a random variable. Define $\mathbb{H} = (\mathcal{H}_t)_{t\geq 0}$ where

$$\mathcal{H}_t = \bigcap_{u>t} \big(\mathcal{F}_u \vee \sigma(\xi)\big)$$

The main theoretical result has been derived by Jacod which we state in the case where ξ takes values in $\mathbb{R}^d$. The conditional probabilities of ξ given $\mathcal{F}_t$ for $t \geq 0$ play a crucial role in this type of filtration expansion.

Assumption 1 (Jacod's criterion) *There exists a σ-finite measure η on $\mathcal{B}(\mathbb{R}^d)$ such that $P(\xi \in \cdot \mid \mathcal{F}_t)(\omega) \ll \eta(\cdot)$ a.s.*

Without loss of generality, η may be chosen as the law of ξ. Under Assumption 1, the $\mathcal{F}_t$-conditional density

$$p_t(u;\omega) = \frac{P(\xi \in du \mid \mathcal{F}_t)(\omega)}{\eta(du)}$$

exists, and can be chosen so that $(u, \omega, t) \mapsto p_t(u;\omega)$ is càdlàg in t and measurable for the optional σ-field associated with the filtration $\widehat{\mathbb{F}}$ given by $\widehat{\mathcal{F}}_t = \cap_{u>t}\mathcal{B}(\mathbb{R}^d) \otimes \mathcal{F}_u$. See Lemma 1.8 in [72].

We provide the explicit decompositions using the following classical result by Jacod, see [72], Theorem 2.5.

Theorem 6 *Let M be an $\mathbb{F}$ local martingale, and assume Assumption 1 is satisfied. Then there exists a set $B \in \mathcal{B}(\mathbb{R}^d)$, with $\eta(B) = 0$, such that*

(i) $\langle p(u), M\rangle$ *exists on* $\{(t,\omega) \mid p_{t-}(u;\omega) > 0\}$ *for every* $u \notin B$,

(ii) there is an increasing predictable process A and an $\widehat{\mathbb{F}}$ predictable function $k_t(u;\omega)$ such that for every $u \notin B$, $\langle p(u), M\rangle_t = \int_0^t k_s(u) p_{s-}(u) dA_s$ on $\{(t,\omega) \mid p_{t-}(u;\omega) > 0\}$,

(iii) $\int_0^t |k_s(\xi)| dA_s < \infty$ a.s. for every $t \geq 0$ and $M_t - \int_0^t k_s(\xi) dA_s$ is an $\mathbb{H}$ local martingale.

We provide now two classical corollaries of Jacod's theorem.

Corollary 1 *If ξ is independent of $\mathbb{F}$ then hypothesis $(H^{'})$ holds between $\mathbb{F}$ and $\mathbb{G}$.*

The next corollary is also known as Jacod's countable expansion theorem.

Corollary 2 *If ξ takes only a countable number of values then hypothesis $(H^{'})$ holds between $\mathbb{F}$ and $\mathbb{G}$.*

Proof. Let $\eta(dx) = \sum_{k=1}^{\infty} P(Z = x_k)\delta_{x_k}(dx)$ be the law of ξ, where $\delta_{x_k}(dx)$ is the Dirac measure at x_k. Then the regular conditional probabilities $P_t(\omega, dx)$ are absolutely continuous w.r.t η with Radon-Nikodym density:

$$\sum_{k=1}^{\infty} \frac{P(\xi = x_k \mid \mathcal{F}_t)}{P(\xi = x_k)} 1_{\{x = x_k\}}$$

The result follows from Theorem 6. ∎

We give now an elementary example of such semimartingale decomposition result.

Example 1 *Take for $\mathbb{F}$ the natural filtration of a Brownian motion B. Let $\xi = B_T$. Then*

$$B_t - \int_0^{t\wedge T} \frac{B_T - B_s}{T - s} ds$$

is an $\mathbb{H}$ Brownian motion where $\mathbb{H} = (\mathcal{H}_t)_{0\leq t<T}$ and $\mathcal{H}_t = \bigcap_{u>t}\left(\mathcal{F}_u \vee \sigma(B_T)\right)$.

Example 1 can be extended to the case $\xi = \int_0^\infty f(s)dB_s$ under good integrability assumptions on f (see [114] and [117]). In this case, conditionnally to $\mathcal{F}_t$, ξ is gaussian with mean $\int_0^t f(s)dB_s$ and variance $\int_0^t f^2(s)ds$. Jeulin and Yor also provided examples where hypothesis $H^{'}$ may not hold but where some local martingales of the base filtration remain semimartingales in the expanded filtration, see [91]. In this paper, they actually caracterise all the martingales of a Brownian filtration $\mathbb{F}$ that remain $\mathbb{H}$ semimartingales, where $\mathbb{H}$ is the initial expansion of $\mathbb{F}$ with B_1, the terminal value of the Brownian motion, and provide their decompositions in $\mathbb{H}$. The same result is available in [117, Chapter V, Theorem 7]. Note that Example 1, where any $\mathbb{F}$ local martingale is claimed to remain a semimartingale in $(\mathcal{H}_t)_{0\leq t<T}$ but not necessarily including T, is consistent with these results. Of course there are also many cases where Jacod's criterion will be violated. Some results when this is the case can be found in [114]. We present in the next subsection the theory of progressive filtration expansion, which was first developed by Yor [127], and further by Jeulin and Yor [92] and Jeulin [89].

1.2.3 Progressive filtration expansions

We will distinguish several cases. First we start with the progressive expansion with an arbitrary random time τ, in which case the results usually hold up to τ. Define the filtration $\mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$ to be the smallest right-continuous filtration containing $\mathbb{F}$ and making τ a stopping time. That is for each $t \geq 0$,

$$\mathcal{G}_t = \bigcap_{u>t} \mathcal{F}_u \vee \sigma(s\wedge\tau, s\leq u)$$

Sometimes, one might need to introduce the larger filtration $\mathbb{G}^\tau = (\mathcal{G}_t^\tau)_{t\geq 0}$ where

$$\mathcal{G}_t^\tau = \Big\{A \in \mathcal{F}_\infty : \exists A_t \in \mathcal{F}_t \mid A\cap\{\tau>t\} = A_t\cap\{\tau>t\}\Big\}$$

We will then focus on the particular case of honest times. This will be useful for our study in section 1.5. In this case, it can be shown that

$$\mathcal{G}_t = \Big\{A \in \mathcal{F}_\infty : \exists A_t, B_t \in \mathcal{F}_t \mid A = (A_t\cap\{\tau>t\})\cup(B_t\cap\{\tau\leq t\})\Big\}$$

We focus then on extending Jeulin-Yor theorem to some more general type of progressive filtration expansions. The main results of this thesis will assume that the times are such that they allow an initial expansion (in the sense that they satisfy Jacod's criterion as in Assumption 1, or more generally that they are such that hypothesis $(H^{'})$ holds between $\mathbb{F}$ and $\mathbb{H}$, the initial expansion of $\mathbb{F}$ with the random time τ).

1.2.3.1 Progressive filtration expansion with an arbitrary random time

Lemma 4 *Let τ be an arbitrary random time. Then* *(i) If H is a $\mathbb{G}^\tau$ predictable process then there exists an $\mathbb{F}$ predictable process J such that $H_t 1_{\{t\leq\tau\}} = J_t 1_{\{t\leq\tau\}}$.*

(ii) Let X be an integrable random variable which is $\mathcal{F}_\infty$ measurable. Then a càdlàg version of the martingale $X_t = E(X \mid \mathcal{G}^\tau_t)$ is given by

$$X_t = \frac{1}{Z_t} 1_{\{\tau>t\}} E(X 1_{\{\tau>t\}} \mid \mathcal{F}_t) + X 1_{\{\tau\leq t\}}$$

Concerning point (ii), note that the quantity in the denominator cannot vanish since $Z_{s-} > 0$ on $\{\tau \geq s\}$. It is also worth mentionning that Z_{s-} is the $\mathbb{F}$ predictable projection of $\mathbf{1}_{\{\tau\geq s\}}$. A proof of these facts can be found in [117]. We recall now the general form of Jeulin-Yor Theorem. Its proof can be found in [90] or [60], for instance.

Theorem 7 (Jeulin-Yor) *Let $\mathbb{F}$ be a filtration, and let $\mathbb{G}$ be the progressive expansion of $\mathbb{F}$ with a nonnegative random variable τ. Define $Z_t = P(\tau > t \mid \mathcal{F}_t)$ and let $Z = M - A$ be its Doob-Meyer decomposition. Then*

$$\mathbf{1}_{\{\tau\leq t\}} - \int_0^{t\wedge\tau} \frac{1}{Z_{s-}} dA_s$$

is a $\mathbb{G}$ martingale. More generally, if H is a bounded $\mathbb{G}^\tau$ predictable process then

$$H_\tau 1_{\{\tau\leq t\}} - \int_0^{t\wedge\tau} \frac{H_s}{Z_{s^-}} dA_s$$

is a $\mathbb{G}^\tau$ martingale.

In [60], this is proved to hold for any progressive expansion $\mathbb{G}$ of $\mathbb{F}$ that satisfies, for each $t \geq 0$,

$$\mathcal{F}_t \cap \{\tau > t\} = \mathcal{G}_t \cap \{\tau > t\}$$

For an arbitrary random time τ, a decomposition formula is available only up to the time τ.

Theorem 8 (Jeulin-Yor) *Fix an $\mathbb{F}$ local martingale M and define $Z_t = P(\tau > t \mid \mathcal{F}_t)$ as the optional projection of $1_{[\![0,\tau[\![}$ onto $\mathbb{F}$, let μ be the martingale part of its Doob-Meyer decomposition, and let J be the dual predictable projection of $\Delta M_\tau \mathbf{1}_{[\![\tau,\infty[\![}$ onto $\mathbb{F}$. Then*

$$M_{t\wedge\tau} - \int_0^{t\wedge\tau} \frac{d\langle M,\mu\rangle_s + dJ_s}{Z_{s-}}$$

is a local martingale in both $\mathbb{G}^\tau$ and $\mathbb{G}$.

1.2.3.2 The case of honest times

The following family of random times, called honest times, is important in progressive filtration expansions.

Definition 2 (Honest times) *Let τ be a random time with values in $[0,\infty]$. τ is said to be honest if for every t, there exists an $\mathcal{F}_t$ measurable random variable τ_t, such that $\tau = \tau_t$ on the set $\{\tau < t\}$.*

These times can be characterized as ends of optional sets.

Lemma 5 *Let τ be a random time. Then τ is an honest time if and only if there exists an optional set H in $[0,\infty] \times \Omega$ such that τ is the end of H.*

Note that ends of optional sets in $[0,\infty[\times\Omega$ do not allow to construct all honest times. The following characterization of $\mathbb{G}$ predictable processes turns out to be very useful.

Lemma 6 *Let τ be an honest time. Then H is a $\mathbb{G}$ predictable process if and only if there exist two $\mathbb{F}$ predictable processes J and K such that $H_t = J_t 1_{\{\tau \geq t\}} + K_t 1_{\{\tau < t\}}$.*

For a honest time τ, a decomposition formula is available not only up to the time τ but on all $\mathbb{R}^+$.

Theorem 9 *An $\mathbb{F}$ local martingale M is a $\mathbb{G}$ semimartingale and*

$$M_t - \int_0^{t\wedge\tau} \frac{d\langle M, Z\rangle_s}{Z_{s^-}} - \int_{t\wedge\tau}^t \frac{d\langle M, Z\rangle_s}{1 - Z_{s^-}}$$

is a $\mathbb{G}$ local martingale. Here Z is the optional projection of $1_{[\![0,\tau[\![}$ onto $\mathbb{F}$.

The previous theorem has been extended by Jeulin to the case of successive progressive expansions with an increasing sequence of honest times. Assume τ^n is an increasing sequence of $\mathbb{F}$ honest times such that $\tau_0 = 0$ and $\sup_n \tau_n = \infty$. Let $\mathbb{G}$ be the smallest filtration containing $\mathbb{F}$ and that makes all $(\tau^n)_{n\geq 1}$ stopping times. Under these assumptions, the following holds (see Jeulin [89, Corollary 5.22]).

Theorem 10 (Jeulin) *Let M be an $\mathbb{F}$ local martingale. Then $M - A$ is a $\mathbb{G}$ local martingale, where*

$$A_t = \sum_{n=0}^{\infty} \int_0^t 1_{\{\tau^n < s \leq \tau^{n+1}\}} \frac{1}{Z^{n+1}_{s^-} - Z^n_{s^-}} d\langle M, M^{n+1} - M^n\rangle_s \tag{1.1}$$

where Z^n is the $\mathbb{F}$ optional projection of $1_{[\![0,\tau^n[\![}$ and M^n is the martingale part in its Doob-Meyer decomposition.

Theorem 10 is a crucial tool to derive some of the results of section 1.5. In the next subsections, we introduce more general types of progressive filtration expansions. In addition to a random time we also expand with an associated random variable corresponding to a *jump size*. We then show how to deal with multiple (even countably many) ordered random times and their associated jump sizes. This allows us to analyze progressive expansions with integrable, finite activity jump processes. Given the work in [60], this turns out to be quite easy to do, but is included first for completeness, and second because it introduces some definitions and tools that would be useful later. The question of interest of this chapter, which is *Does a given semimartingale of the base filtration remain a semimartingale* *in the expansion of this filtration with a given process?* is naturally postponed to the next sections of the chapter.

1.2.3.3 The (τ, X)-progressive expansion

We present and study in some detail a first generalization of the minimal progressive expansion.

Definition 3 *Given a filtration $\mathbb{F}$, a random time τ and a random variable X, let $\mathbb{G}$ be the smallest right-continuous filtration containing $\mathbb{F}$, that makes τ a stopping time and such that X is $\mathcal{G}_\tau$-measurable. We refer to $\mathbb{G}$ as the* (τ, X)-progressive expansion *of $\mathbb{F}$.*

Note that $\mathbb{G}$ is well-defined, since it is an intersection over a non-empty set of filtrations (the constant filtration $\mathbb{H}$ given by $\mathcal{H}_t = \mathcal{F}$ for all t is among the candidates.) When X is constant, the (τ, X)-progressive expansion coincides with the usual progressive expansion with τ. Our (τ, X)-progressive expansion can be seen as a particular case of a more general progressive filtration expansion introduced by the French school where some given σ-field $\mathcal{E}$ is added to the base filtration at time τ. In our case, the σ-field $\mathcal{E}$ is $\sigma(X)$. This section also uses some ideas already present in Dellacherie-Meyer's work (see for instance [35] and [37]). One can easily obtain an explicit description of the (τ, X)-progressively expanded filtration. Let $\mathbb{F}^\tau$ be the progressive expansion of $\mathbb{F}$ with τ.

Lemma 7 *The (τ, X)-progressive expansion $\mathbb{G}$ of $\mathbb{F}$ is given by $\mathcal{G}_t = \bigcap_{u>t} \mathcal{G}_u^0$ where*

$$\mathcal{G}_t^0 = \mathcal{F}_t^\tau \vee \sigma(\mathbf{1}_{\{\tau \le t\}} X).$$

Proof. Denote the right side by $\mathcal{H}_t^0$, that is $\mathcal{H}_t^0 = \mathcal{F}_t^\tau \vee \sigma(\mathbf{1}_{\{\tau \le t\}} X)$. The inclusion $\mathcal{H}_t^0 \subset \mathcal{G}_t$ follows from the Monotone Class Theorem if we can show that $Y_t h(\mathbf{1}_{\{\tau \le t\}} X)$ is $\mathcal{G}_t$-measurable for every bounded measurable function $h : \mathbb{R} \to \mathbb{R}$ and every bounded $\mathcal{F}_t^\tau$-measurable random variable Y_t. Since $\mathcal{F}_t^\tau \subset \mathcal{G}_t$ it suffices to prove that $\mathbf{1}_{\{\tau \le t\}} X$ is $\mathcal{G}_t$-measurable. For an arbitrary set $C \in \mathcal{B}(\mathbb{R})$,

$$\{\mathbf{1}_{\{\tau \le t\}} X \in C\} = (\{X \in C\} \cap \{\tau \le t\}) \cup (\{0 \in C\} \cap \{\tau > t\}).$$

Now, by definition of $\mathbb{G}$, $\{X \in C\} \in \mathcal{G}_\tau = \{A \in \mathcal{G}_\infty : A \cap \{\tau \le t\} \in \mathcal{G}_t, \forall t \ge 0\}$, so $\{X \in C\} \cap \{\tau \le t\} \in \mathcal{G}_t$. The right side of the above display is thus $\mathcal{G}_t$-measurable, as required. Take now intersections to conclude for the inclusion of the right-continuous modifications.

For $\mathcal{G}_t \subset \mathcal{H}_t = \bigcap_{u>t} \mathcal{H}_u^0$, first note that $\mathcal{F}_t \subset \mathcal{H}_t$, that τ is an $\mathbb{H}$ stopping time and that $\mathbb{H}$ is right-continuous. Furthermore, $X = \mathbf{1}_{\{\tau \le \infty\}} X$ is $\mathcal{H}_\infty$-measurable, and

$$\{X \in C\} \cap \{\tau \le s\} = \{\mathbf{1}_{\{\tau \le s\}} X \in C\} \cap \{\tau \le s\} \in \mathcal{H}_s$$

holds for every $s \ge 0$, so X is $\mathcal{H}_\tau$-measurable. Since $\mathbb{G}$ is the smallest filtration with these properties, the result follows. ■

In the same spirit as of the work by Guo and Zeng (see [60]), the following lemma is the crucial result that enables us to extend the Jeulin-Yor Theorem in a straightforward way.

Lemma 8 *Let $\mathbb{G}$ be the (τ, X)-progressive expansion of $\mathbb{F}$. Then*

$$\{\tau > t\} \cap \mathcal{G}_t = \{\tau > t\} \cap \mathcal{F}_t$$

for all $t \geq 0$.

Proof. It is well known that $\{\tau > t\} \cap \mathcal{F}_t^\tau = \{\tau > t\} \cap \mathcal{F}_t$. With the representation from Lemma 7 in mind, let $H_t = Y_t h(\mathbf{1}_{\{\tau \leq t\}} X)$, where Y_t is $\mathcal{F}_t^\tau$-measurable and bounded, and h is a bounded Borel function. Then $H_t \mathbf{1}_{\{\tau > t\}} = Y_t h(0) \mathbf{1}_{\{\tau > t\}}$, which is measurable with respect to $\{\tau > t\} \cap \mathcal{F}_t^\tau = \{\tau > t\} \cap \mathcal{F}_t$. The Monotone Class Theorem now proves that $\{\tau > t\} \cap \mathcal{G}_t \subset \{\tau > t\} \cap \mathcal{F}_t$. The reverse inclusion is clear. ■

We can now compute the $\mathbb{G}$ compensator of $X\mathbf{1}_{\{\tau \leq t\}}$. The next result is completely analogous to the Jeulin-Yor Theorem, with an almost identical proof. The proof we give here follows that of Guo and Zeng [60].

Theorem 11 *Let X be integrable, and suppose $\mathbb{G}$ is a filtration that contains $\mathbb{F}$, makes τ a stopping time, and is such that X is $\mathcal{G}_\tau$-measurable. Assume also that $\{\tau > t\} \cap \mathcal{G}_t = \{\tau > t\} \cap \mathcal{F}_t$ for every $t \geq 0$. Let $Z_t = P(\tau > t \mid \mathcal{F}_t)$ and let A^X be the $\mathbb{F}$ dual predictable projection of $X\mathbf{1}_{\{\tau \leq t\}}$. Then the process*

$$X\mathbf{1}_{\{\tau \leq t\}} - \int_0^{t \wedge \tau} \frac{1}{Z_{s-}} dA_s^X$$

is a $\mathbb{G}$ martingale.

Proof. The result follows if we prove that for any bounded $\mathbb{G}$ predictable process H,

$$E\left\{\int_0^\infty H_s d(X\mathbf{1}_{\{\tau \leq \cdot\}})_s\right\} = E\left\{\int_0^\infty H_s \frac{1}{Z_{s-}} \mathbf{1}_{\{\tau \geq s\}} dA_s^X\right\}.$$

Let H be such a process. Then $\{\tau > t\} \cap \mathcal{G}_t = \{\tau > t\} \cap \mathcal{F}_t$ for every $t \geq 0$ implies that there exists an $\mathbb{F}$ predictable bounded process J, such that $H_t = J_t$ on the set $\{0 \leq t \leq \tau\}$; for a proof of this fact, see e.g. [60]. Then

$$\begin{aligned} E\left\{\int_0^\infty H_s d(X\mathbf{1}_{\{\tau \leq \cdot\}})_s\right\} &= E(XH_\tau \mathbf{1}_{\{\tau < \infty\}}) \\ &= E(XJ_\tau \mathbf{1}_{\{\tau < \infty\}}) = E\left\{\int_0^\infty J_s d(X\mathbf{1}_{\{\tau \leq \cdot\}})_s\right\}. \end{aligned}$$

By definition of the dual predictable projection, the right side is equal to $E(\int_0^\infty J_s dA_s^X)$. Moreover,

$$E(\int_0^\infty J_s dA_s^X) = E(\int_0^\infty \frac{J_s}{Z_{s-}} Z_{s-} dA_s^X) = E(\int_0^\infty \frac{J_s}{Z_{s-}} \mathbf{1}_{\{\tau \geq s\}} dA_s^X),$$

where the second equality follows from the fact that Z_{s-} is the predictable projection of $\mathbf{1}_{\{\tau \geq s\}}$ (a proof of this fact as well as the fact that $Z_{s-} > 0$ on $\{\tau \geq s\}$ can be found in [117]). The right side above is equal to $E(\int_0^\infty H_s \frac{1}{Z_{s-}} \mathbf{1}_{\{\tau \geq s\}} dA_s^X)$, which ends the proof. ■

Remark 1 *First, Theorem 11 reduces to the original Jeulin Yor Theorem (Theorem 7) when $X = 1$. Second, when $\mathbb{G}$ is the (τ, X)-progressive expansion of $\mathbb{F}$, Lemma 8 implies that Theorem 11 is applicable. Notice also that Theorem 7 still gives the correct compensator of $\mathbf{1}_{\{\tau \leq t\}}$ in the filtration $\mathbb{G}$, precisely because the criterion $\mathcal{G}_t \cap \{\tau > t\} = \mathcal{F}_t \cap \{\tau > t\}$ is satisfied. See [60] where the authors were the first to point out this remark.*

Remark 2 *In credit risk applications, X could represent the recovery in case default occurs before some specified time horizon. It may be reasonable to assume that this amount is not observable in the base filtration $\mathbb{F}$. Theorem 11 allows us to model both the default time and the recovery value independently of $\mathbb{F}$, and then obtain pricing formulas in terms of $\mathbb{F}$-conditional quantities.*

1.2.3.4 Expansion with a jump process

We now have the machinery to expand the base filtration $\mathbb{F}$ with a jump process of the form $Y_t = \sum_{i=1}^\infty X_i \mathbf{1}_{\{\tau_i \leq t\}}$, where each X_i is integrable, and the jump times τ_i form a strictly increasing sequence tending to infinity almost surely. The expanded filtration is the smallest filtration containing $\mathbb{F}$ and for which Y is adapted. The canonical example of such a process Y is the compound Poisson process, but any finite activity pure jump process with integrable jumps fits into this framework. We will carry out the analysis under somewhat more general assumptions on the expanded filtrations. The purpose is to highlight exactly which properties are relevant. Consider the sequence of ranked times $0 = \tau_0 \leq \tau_1 \leq \tau_2 \leq \dots$ together with a sequence of integrable jump sizes $X_1, X_2, \dots$ (conventionally, $X_0 = 0$.) Suppose also we have an increasing sequence of filtrations $\mathbb{F} = \mathbb{F}^0 \subset \mathbb{F}^1 \subset \mathbb{F}^2 \subset \dots$ with the property that τ_i is an $\mathbb{F}^i$ stopping time and X_i is $\mathcal{F}^i_{\tau_i}$ measurable, for each i. Let $\mathbb{G}$ be the filtration generated by all the $\mathbb{F}^i$. We impose the following condition on the filtrations $\mathbb{F}^i$:

Condition 1 *For every $i \geq 1$ and every t, $\{t < \tau_i\} \cap \mathcal{F}^{i-1}_t = \{t < \tau_i\} \cap \mathcal{F}^i_t$.*

For example and as proved in Lemma 8, this condition is satisfied if $\mathbb{F}^i$ is the (τ_i, X_i)-progressive expansion of $\mathbb{F}^{i-1}$. A consequence is the following lemma, which is what allows us to apply Theorem 11 in this setting.

Lemma 9 *If Condition 1 holds, then for every $j \geq i \geq 1$ we have, for every t,*

$$\{t < \tau_i\} \cap \mathcal{F}^{i-1}_t = \{t < \tau_i\} \cap \mathcal{F}^j_t.$$

As a consequence, we also have

$$\{t < \tau_i\} \cap \mathcal{F}^{i-1}_t = \{t < \tau_i\} \cap \mathcal{G}_t.$$

Proof. We start with the first statement. By Condition 1, $\{t < \tau_{j+1}\} \cap \mathcal{F}_t^j = \{t < \tau_{j+1}\} \cap \mathcal{F}_t^{j+1}$. We intersect both side by $\{t < \tau_i\}$ and use that $\tau_i \le \tau_{j+1}$ implies $\{t < \tau_i\} \subset \{t < \tau_{j+1}\}$ to get $\{t < \tau_i\} \cap \mathcal{F}_t^j = \{t < \tau_i\} \cap \mathcal{F}_t^{j+1}$. Hence if the statement is true for j, it is also true for $j+1$. It is true for $j = i$ by assumption, so the result follows by induction. Regarding the second statement, since $\mathcal{G}_t$ is generated by all the $\mathcal{F}_t^j$, the result follows from the first one using the Monotone Class Theorem. ■

Next we introduce some notation.

- $Z_t^i = P(\tau_i > t \mid \mathcal{F}_t^{i-1})$, the optional projection of $\mathbf{1}_{\{\tau_i > \cdot\}}$ onto $\mathbb{F}^{i-1}$.
- A_t^i, the $\mathbb{F}^{i-1}$ dual predictable projection of $X_i \mathbf{1}_{\{\tau_i \le \cdot\}}$.
- $\Lambda_t^i = \int_0^{t\wedge\tau_i} \frac{1}{Z_{s-}^i} dA_s^i$.

The next result is immediate.

Lemma 10 *Suppose Condition 1 holds. Then for every $i \ge 1$, the process*

$$M_t^i = X_i \mathbf{1}_{\{\tau_i \le t\}} - \Lambda_t^i$$

is a $\mathbb{G}$ martingale.

Proof. This follows directly from Lemma 9 and Theorem 11. ■

Lemma 10 suffices if we are interested in the times τ_i separately. To deal with the entire counting process that they generate, given by

$$N_t = \sum_{i\ge 1} X_i \mathbf{1}_{\{\tau_i \le t\}},$$

some condition is needed to ensure that the process does not blow up, and that the jumps are not too large (there could be large jumps if several of the τ_i coincide.) One possible condition is the following.

Condition 2 *The sequence (τ_i) is strictly increasing to infinity. That is, $\tau_i > \tau_{i-1}$ on the set where $\tau_{i-1} < \infty$, and $\lim_{i\to\infty} \tau_i = \infty$ a.s.*

This certainly guarantees that N_t is a.s. finite for all t and that the i:th jump has size X_i. Furthermore, we get the following result, whose proof is standard.

Lemma 11 *Under Condition 2, dA_t^i does not charge the set $\{t : t \le \tau_{i-1}\}$.*

Proof. We provide two proofs. By basic properties of dual predictable projections, for each $i \ge 1$ we have

$$E\left[\int_0^\infty \mathbf{1}_{\{t\le\tau_{i-1}\}} dA_t^i\right] = E\left[\int_0^\infty \mathbf{1}_{\{t\le\tau_{i-1}\}} d(X_i \mathbf{1}_{\{\tau_i\le\cdot\}})_t\right] = E\left[X_i \mathbf{1}_{\{\tau_i\le\tau_{i-1}\}} \mathbf{1}_{\{\tau_i<\infty\}}\right],$$

which is zero by Condition 2. We give now an alternative proof of this result. Let $Z_t^i = P(\tau_i > t | \mathcal{F}_t^{i-1}) = M_t^i - A_t^i$ be the Doob-Meyer decomposition of the supermartingale Z^i.

By Condition 2, $Z^i_{\tau_{i-1}} = P(\tau_i > \tau_{i-1} | F^{i-1}_{\tau_{i-1}}) = 1$, so by the supermartingale property we have, for all $t \leq \tau_{i-1}$, that $1 \geq Z^i_t \geq E(Z^i_{\tau_{i-1}} | \mathcal{F}^{i-1}_t) = 1$. Hence $Z^i_t = 1$ on $\{t \leq \tau_{i-1}\}$. Thus on $\{t \leq \tau_{i-1}\}$, $A^i_t = M^i_t - 1$ is an $\mathbb{F}^{i-1}$ predictable increasing process and a uniformly integrable martingale, so it is constant there. This proves again the lemma. ■

We can now extend the Jeulin-Yor Theorem to non-explosive counting processes with non-overlapping jumps.

Theorem 12 *Suppose Conditions 1 and 2 hold. Then the process*

$$\Lambda_t = \sum_{i \geq 1} \Lambda^i_t$$

is well-defined and finite a.s., and $N_t - \Lambda_t$ is a $\mathbb{G}$ local martingale.

Proof. Lemma 11 says that dA^i_t does not charge $\{t : t \leq \tau_{i-1}\}$, and so

$$\Lambda^i_t = \int_0^t \mathbf{1}_{\{\tau_{i-1} < s \leq \tau_i\}} \frac{1}{Z^i_{s-}} dA^i_s.$$

Thus $\Lambda^i_t = 0$ for $t \leq \tau_{i-1}$, and since $\tau_i \to \infty$ as $i \to \infty$, the sum $\sum_{i \geq 1} \Lambda^i_t$ contains finitely many terms, each of which is finite. So Λ_t is well-defined and finite.

Next, since $\tau_i > \tau_{i-1}$ a.s. and since $\Lambda^i_t = 0$ for $t \leq \tau_{i-1}$,

$$N_{t \wedge \tau_n} - \Lambda_{t \wedge \tau_n} = \sum_{i=1}^n \left(\mathbf{1}_{\{t \leq \tau_i\}} - \Lambda^i_t \right).$$

This is a martingale by Lemma 10. Since $\tau_n \to \infty$, this shows that $N - \Lambda$ is a local martingale with reducing sequence $(\tau_n)_{n \geq 1}$. ■

If it turned out to be quite easy to obtain the results for the compensator of the jump process in the expanded filtration $\mathbb{G}$, it is much harder to see if an $\mathbb{F}$ local martingale remains a $\mathbb{G}$ semimartingale. This is the kind of questions we deal with in sections 1.4 and 1.5 where we expand the base filtration $\mathbb{F}$ with a process. Note also that the processes we will expand with are much more general than counting processes. However, in order to derive these resuts, we need to introduce the concepts of weak convergence of σ-fields and of filtrations.

1.2.4 Weak convergence of filtrations and of σ-fields

Hoover [68], following remarks by M. Barlow and S. Jacka, introduced the weak convergence of σ-fields and of filtrations in 1991. The next big step was in 2000 with the seminal paper of Antonelli and Kohatsu-Higa [7]. This was quickly followed by the work of Coquet, Mémin and Mackevicius [29] and by Coquet, Mémin and Slominsky [28]. We will recall fundamental results on the topic but we refer the interested reader to [28] and [29] for details. In these papers, all filtrations are indexed by a compact time interval $[0, T]$ and we work also within a finite time horizon framework and assume that a probability space $(\Omega, \mathcal{H}, P)$ and a positive integer T are given. All filtrations considered are assumed to be completed by the P-null sets of $\mathcal{H}$. By the natural filtration of a process X, we mean the right-continuous filtration associated with the natural filtration of X. The concepts of weak convergence of σ-fields and of filtrations rely on the topology imposed on the space of càdlàg processes and we use the Skorohod J_1 topology as it is done in [28]. An outline of this section is the following. In subsection 1.2.4.1, we recall basic facts on the weak convergence of σ-fields and establish fundamental lemmas for subsequent use. A crucial lemma is given in subsection 1.2.4.2, and allows to approximate a given $\mathbb{G}$ stopping time τ with a sequence of $\mathbb{G}^n$ stopping times τ^n given that the σ-fields $\mathcal{G}^n_t$ converge weakly to $\mathcal{G}_t$ for each t. The last subsection provides a sufficient condition for the semimartingale property to hold for a given càdlàg adapted process based on the weak convergence of σ-fields. The sufficient condition we provide at this point is unlikely to hold in a filtration expansion context, however the proof of this result underlines what can go wrong under the more natural assumptions considered in the sections 1.4 and 1.5.

1.2.4.1 Definitions and fundamental results

Let $\mathbb{D}$ be the space of càdlàg functions from $[0, T]$ into $\mathbb{R}$. Let Λ be the set of time changes from $[0, T]$ into $[0, T]$, i.e. the set of all continuous strictly increasing functions $\lambda : [0, T] \to [0, T]$ such that $\lambda(0) = 0$ and $\lambda(T) = T$. We define the Skorohod distance as follows

$$d_S(x, y) = \inf_{\lambda \in \Lambda} \left\{ ||\lambda - Id||_\infty \vee ||x - y \circ \lambda||_\infty \right\}$$

for each x and y in $\mathbb{D}$. Let $(X^n)_{n \geq 1}$ and X be càdlàg processes (i.e. whose paths are in $\mathbb{D}$), indexed by $[0, T]$ and defined on $(\Omega, \mathcal{H}, P)$. We will write $X^n \xrightarrow{P} X$ when $(X^n)_{n \geq 1}$ converges in probability under the Skorohod J_1 topology to X i.e. when the sequence of random variables $(d_S(X^n, X))_{n \geq 1}$ converges in probability to zero. We can now introduce the concepts of weak convergence of σ-fields and of filtrations.

Definition 4 *A sequence of σ-algebras $\mathcal{A}^n$ converges weakly to a σ-algebra $\mathcal{A}$ if and only if for all $B \in \mathcal{A}$, $E(1_B \mid \mathcal{A}^n)$ converges in probability to 1_B. We write $\mathcal{A}^n \xrightarrow{w} \mathcal{A}$.*

Definition 5 *A sequence of right-continuous filtrations $\mathbb{F}^n$ converges weakly to a filtration $\mathbb{F}$ if and only if for all $B \in \mathcal{F}_T$, the sequence of càdlàg martingales $E(1_B \mid \mathcal{F}^n_.)$ converges in probability under the Skorohod J_1 topology on $\mathbb{D}$ to the martingale $E(1_B \mid \mathcal{F}_.)$. We write $\mathbb{F}^n \xrightarrow{w} \mathbb{F}$.*

The following lemmas provide characterizations of the weak convergence of σ-fields and filtrations. We refer to [28] for the proofs.

Lemma 12 *A sequence of σ-algebras $\mathcal{A}^n$ converges weakly to a σ-algebra $\mathcal{A}$ if and only if $E(Z \mid \mathcal{A}^n)$ converges in probability to Z for any integrable and $\mathcal{A}$ measurable random variable Z.*

Lemma 13 *A sequence of filtrations* $\mathbb{F}^n$ *converges weakly to a filtration* $\mathbb{F}$ *if and only if* $E(Z \mid \mathcal{F}^n_.)$ *converges in probability under the Skorohod* J_1 *topology to* $E(Z \mid \mathcal{F}_.)$, *for any integrable,* $\mathcal{F}_T$ *measurable random variable* Z.

The weak convergence of the σ-fields $\mathcal{F}^n_t$ to $\mathcal{F}_t$ for all t does not imply the weak convergence of the filtrations $\mathbb{F}^n$ to $\mathbb{F}$. The reverse implication does not hold neither. Coquet, Mémin and Slominsky provide a characterization of weak convergence of filtrations when the limiting filtration is the natural filtration of some càdlàg process X. The result, recalled below, is Lemma 3 in [28].

Lemma 14 *Let* $(X^n)_{n\geq 1}$ *be a sequence of càdlàg processes converging in probability to a càdlàg process* X. *Let* $\mathbb{F}^n$ *and* $\mathbb{F}$ *be the natural filtrations of* X^n *and* X *respectively. The following conditions are equivalent.*

(i) $\mathbb{F}^n \xrightarrow{w} \mathbb{F}$

(ii) $E(f(X_{t_1},\dots,X_{t_k}) \mid \mathcal{F}^n_.)$ *converges in probability to* $E(f(X_{t_1},\dots,X_{t_k}) \mid \mathcal{F}_.)$ *for every bounded continuous function* $f : \mathbb{R}^k \to \mathbb{R}$ *and every points* $t_1,\dots,t_k$ *such that for each* $1 \leq i \leq k$, $P(\Delta X_{t_i} \neq 0) = 0$ *or* $t_i = T$, *for each* $k \geq 1$.

We provide a similar result for the weak convergence of σ-fields when the limiting σ-field is generated by some càdlàg process X.

Lemma 15 *Let* X *be a càdlàg process. Define* $\mathcal{A} = \sigma(X_t, 0 \leq t \leq T)$ *and let* $(\mathcal{A}^n)_{n\geq 1}$ *be a sequence of* σ*-fields. Then* $\mathcal{A}^n \xrightarrow{w} \mathcal{A}$ *if and only if*

$$E(f(X_{t_1},\dots,X_{t_k}) \mid \mathcal{A}^n) \xrightarrow{P} f(X_{t_1},\dots,X_{t_k})$$

for all $k \in \mathbb{N}$, $t_1,\dots,t_k$ *points of a dense subset* $\mathcal{D}$ *of* $[0,T]$ *containing* T *and for any continuous and bounded function* $f : \mathbb{R}^k \to \mathbb{R}$.

Proof. Necessity follows from the definition of the weak convergence of σ-fields. Let us prove the sufficiency. Let $A \in \mathcal{A}$ and $\varepsilon > 0$. There exist $k \in \mathbb{N}$ and $t_1,\dots,t_k$ in $\mathcal{D}$ such that

$$E(|f(X_{t_1},\dots,X_{t_k}) - 1_A|) < \varepsilon.$$

Let $\eta > 0$. We need to show that $P(|E(1_A \mid \mathcal{A}^n) - 1_A| \geq \eta)$ converges to zero.

$$\begin{aligned}
&P(|E(1_A \mid \mathcal{A}^n) - 1_A| \geq \eta) \leq P(|E(1_A \mid \mathcal{A}^n) - E(f(X_{t_1},\dots,X_{t_k}) \mid \mathcal{A}^n)| \geq \frac{\eta}{3}) \\
&\quad + P(|E(f(X_{t_1},\dots,X_{t_k}) \mid \mathcal{A}^n) - f(X_{t_1},\dots,X_{t_k})| \geq \frac{\eta}{3}) + P(|f(X_{t_1},\dots,X_{t_k}) - 1_A| \geq \frac{\eta}{3}) \\
&\quad \leq \frac{6}{\eta} E(|f(X_{t_1},\dots,X_{t_k}) - 1_A|) + P(|E(f(X_{t_1},\dots,X_{t_k}) \mid \mathcal{A}^n) - f(X_{t_1},\dots,X_{t_k})| \geq \frac{\eta}{3}) \\
&\quad \leq \frac{6}{\eta}\varepsilon + P(|E(f(X_{t_1},\dots,X_{t_k}) \mid \mathcal{A}^n) - f(X_{t_1},\dots,X_{t_k})| \geq \frac{\eta}{3})
\end{aligned}$$

where the second inequality follows from the Markov inequality. By assumption, there exists N such that for all $n \geq N$,

$$P(|E(f(X_{t_1},\dots,X_{t_k}) \mid \mathcal{A}^n) - f(X_{t_1},\dots,X_{t_k})| \geq \frac{\eta}{3}) \leq \varepsilon$$

hence $P(|E(1_A \mid \mathcal{A}^n) - 1_A| \geq \eta) \leq (\frac{6}{\eta} + 1)\varepsilon$. ∎

In [28], the authors provide cases where the weak convergence of a sequence of natural filtrations of given càdlàg processes is guaranteed. We provide here a similar result for the weak convergence of the associated σ-fields.

Lemma 16 *Let $(X^n)_{n\geq 1}$ be a sequence of càdlàg processes converging in probability to a càdlàg process X. Let $\mathbb{F}^n$ and $\mathbb{F}$ be the natural filtrations of X^n and X respectively. Then $\mathcal{F}_t^n \xrightarrow{w} \mathcal{F}_t$ for all t such that $P(\Delta X_t \neq 0) = 0$.*

Proof. Let t be such that $P(\Delta X_t \neq 0) = 0$. Since X is càdlàg, there exists a dense subset $\mathcal{D}$ of $[0, t]$ such that for each $k \in \mathbb{N}$, and $t_1, \ldots, t_k \leq t$ in $\mathcal{D}$, $P(\Delta X_{t_i} \neq 0) = 0$, for all $1 \leq i \leq k$. By assumption $t \in \mathcal{D}$. Let $f : \mathbb{R}^k \to \mathbb{R}$ be a continuous and bounded function. By Lemma 15, it suffices to show that

$$E(f(X_{t_1}, \ldots, X_{t_k}) \mid \mathcal{F}_t^n) \xrightarrow{P} f(X_{t_1}, \ldots, X_{t_k})$$

An application of Markov's inequality leads to the following estimate

$$\begin{aligned}
P(|E(f(X_{t_1}, \ldots, &X_{t_k}) \mid \mathcal{F}_t^n) - (X_{t_1}, \ldots, X_{t_k})| \geq \eta) \\
&\leq P(|E(f(X_{t_1}, \ldots, X_{t_k}) - f(X_{t_1}^n, \ldots, X_{t_k}^n) \mid \mathcal{F}_t^n)| \geq \frac{\eta}{2}) \\
&+ P(|E(f(X_{t_1}^n, \ldots, X_{t_k}^n) \mid \mathcal{F}_t^n) - f(X_{t_1}, \ldots, X_{t_k})| \geq \frac{\eta}{2}) \\
&\leq \frac{4}{\eta} E(|f(X_{t_1}^n, \ldots, X_{t_k}^n) - f(X_{t_1}, \ldots, X_{t_k})|)
\end{aligned}$$

Since $X^n \xrightarrow{P} X$ and $P(\Delta X_{t_i} \neq 0) = 0$, for all $1 \leq i \leq k$, it follows that

$$(X_{t_1}^n, \ldots X_{t_k}^n) \xrightarrow{P} (X_{t_1}, \ldots X_{t_k}) \tag{1.2}$$

and hence $f(X_{t_1}^n, \ldots X_{t_k}^n)$ converges in L^1 to $f(X_{t_1}, \ldots X_{t_k})$. This ends the proof of the lemma. ∎

Definition 6 (Fixed time of discontinuity) *For a given càdlàg process X, a time t such that $P(\Delta X_t \neq 0) > 0$ will be called a fixed time of discontinuity of X. We will say that X has no fixed times of discontinuity if $P(\Delta X_t \neq 0) = 0$ for all $0 \leq t \leq T$.*

Lemma 16 can be improved when the sequence X^n is the discretization of the càdlàg process X along some refining sequence of subdivisions $(\pi_n)_{n\geq 1}$ such that each fixed time of discontinuity of X belongs to $\cap_n \pi_n$.

Lemma 17 *Let X be a càdlàg process. Consider a sequence of subdivisions $(\pi_n = \{t_k^n\}, n \geq 1)$ whose mesh tends to zero and let X_n be the discretized process defined by $X_t^n = X_{t_k^n}$, for all $t_n^k \leq t < t_{k+1}^n$. Let $\mathbb{F}$ and $\mathbb{F}^n$ be the natural filtrations of X and X^n. If each fixed time of discontinuity of X belongs to $\cap_n \pi_n$, then $\mathcal{F}_t^n \xrightarrow{w} \mathcal{F}_t$, for all t.*

Proof. The proof is essentially the same as that of Lemma 16. Now, equation (1.2) holds because the subdivision contains the discontinuity points of X. ■

We will also need the two following lemmas from the theory of weak convergence of σ-fields. The first result is proved in [29] and the second one in [28].

Lemma 18 *Let $(\mathcal{A}^n)_{n\geq 1}$ and $(\mathcal{B}^n)_{n\geq 1}$ be two sequences of σ-fields that weakly converge to $\mathcal{A}$ and $\mathcal{B}$, respectively. Then*

$$\mathcal{A}^n \vee \mathcal{B}^n \xrightarrow{w} \mathcal{A} \vee \mathcal{B}$$

Lemma 19 *Let $(\mathcal{A}^n)_{n\geq 1}$ and $(\mathcal{B}^n)_{n\geq 1}$ be two sequences of σ-fields such that $\mathcal{A}^n \subset \mathcal{B}^n$ for all n. Let $\mathcal{A}$ be a σ-field. If $\mathcal{A}^n \xrightarrow{w} \mathcal{A}$ then $\mathcal{B}^n \xrightarrow{w} \mathcal{A}$.*

As pointed out in [28], the results in Lemmas 18 and 19 are not true as far as one is interested in weak convergence of filtrations.

1.2.4.2 Approximation of a given stopping time

Let $(\mathbb{G}^n)_{n\geq 1}$ be a sequence of right-continuous filtrations and let $\mathbb{G}$ be a right-continuous filtration such that $\mathcal{G}_t^n \xrightarrow{w} \mathcal{G}_t$ for all t. In order to obtain our filtration expansion results, we need a key theorem that guarantees the $\mathbb{G}$ semimartingale property of a limit of $\mathbb{G}^n$ semimartingales as in Theorem 22. The following lemma, which permits to approximate any $\mathbb{G}$ bounded stopping time τ by a sequence of $\mathbb{G}^n$ stopping times, will be of crucial importance in the proof of Theorem 22, Part (ii). We prove this result using successive approximations in the case where τ takes a finite number of values and show how this property is inherited by bounded stopping times. We do not study the general case (unbounded stopping times) since we are working on the finite time interval $[0, T]$.

Lemma 20 *Let $(\mathbb{G}^n)_{n\geq 1}$ be a sequence of right-continuous filtrations and let $\mathbb{G}$ be a right-continuous filtration such that $\mathcal{G}_t^n \xrightarrow{w} \mathcal{G}_t$ for all t. Let τ be a bounded $\mathbb{G}$ stopping time. Then there exists $\phi : \mathbb{N} \to \mathbb{N}$ strictly increasing and a bounded sequence $(\tau_n)_{n\geq 1}$ such that the subsequence $(\tau_{\phi(n)})_{n\geq 1}$ converges in probability to τ and each $\tau_{\phi(n)}$ is a $\mathbb{G}^{\phi(n)}$ stopping time.*

Proof. Let τ be a $\mathbb{G}$ stopping time bounded by T. Then there exists a sequence τ_n of $\mathbb{G}$ stopping times decreasing a.s. to τ and taking values in $\left\{\frac{k}{2^n}, k \in \{0, 1, \cdots, [2^n T] + 1\}\right\}$. This is true since the sequence $\tau_n = \frac{[2^n \tau]+1}{2^n}$ obviously works. Hence τ_n takes a finite number of values. We claim that

Claim. *for each n, we can construct a sequence $(\tau_{n,m})_{m\geq 1}$ converging in probability to τ_n, and such that $\tau_{n,m}$ is a $\mathbb{G}^m$ stopping time, for each m.*

Assume we can do so and let $\eta > 0$ and $\varepsilon > 0$. Then for each n, $\lim_{m\to\infty} P(|\tau_{n,m} - \tau_n| > \frac{\eta}{2}) = 0$, i.e. for each n there exists M_n such that for all $m \geq M_n$, $P(|\tau_{n,m} - \tau_n| > \frac{\eta}{2}) \leq \frac{\varepsilon}{2}$. Define $\phi(1) = M_1$ and $\phi(n) = \max(M_n, \phi(n-1) + 1)$ by induction. The application

$\phi : \mathbb{N} \to \mathbb{N}$ is strictly increasing, and for each n, $\tau_{n,\phi(n)}$ is a $\mathbb{G}^{\phi(n)}$ stopping time and $P(|\tau_{n,\phi(n)} - \tau_n| > \frac{\eta}{2}) \leq \frac{\varepsilon}{2}$. It follows that

$$P(|\tau_{n,\phi(n)} - \tau| > \eta) \leq P(|\tau_{n,\phi(n)} - \tau_n| > \frac{\eta}{2}) + P(|\tau_n - \tau| > \frac{\eta}{2}) \leq \frac{\varepsilon}{2} + P(|\tau_n - \tau| > \frac{\eta}{2})$$

Since τ_n converges to τ, there exists some n_0, such that for all $n \geq n_0$, $P(|\tau_n - \tau| > \frac{\eta}{2}) \leq \frac{\varepsilon}{2}$. Hence $\tau_{n,\phi(n)} \xrightarrow{P} \tau$. So in order to prove the lemma, it only remains to prove the claim above.

Proof of the claim. We drop the index n and assume that τ is a $\mathbb{G}$ stopping time that takes a finite number of values $t_1, \cdots, t_M$. Since $\mathbb{G}$ is right-continuous, $1_{\{\tau=t_i\}}$ is $\mathcal{G}_{t_i}$ measurable, and since by assumption, for all i, $\mathcal{G}^m_{t_i} \xrightarrow{w} \mathcal{G}_{t_i}$, it follows that for all i

$$E(1_{\{\tau=t_i\}} \mid \mathcal{G}^m_{t_i}) \xrightarrow{P} 1_{\{\tau=t_i\}}$$

Now for $i = 1$, we can extract a subsequence $E(1_{\{\tau=t_1\}} \mid \mathcal{G}^{\phi_1(m)}_{t_1})$ converging to $1_{\{\tau=t_1\}}$ a.s. and any sub-subsequence will also converge to $1_{\{\tau=t_1\}}$ a.s. Also, $\mathcal{G}^{\phi_1(m)}_{t_2} \xrightarrow{w} \mathcal{G}_{t_2}$, hence $E(1_{\{\tau=t_2\}} \mid \mathcal{G}^{\phi_1(m)}_{t_2}) \xrightarrow{P} 1_{\{\tau=t_2\}}$, and we can extract a further subsequence $E(1_{\{\tau=t_2\}} \mid \mathcal{G}^{\phi_1(\phi_2(m))}_{t_2})$ that converges a.s. to $1_{\{\tau=t_2\}}$. Since we have a finite number of possible values, we can repeat this reasoning up to time t_M. Define then $\phi = \phi_1 \circ \phi_2 \circ \cdots \circ \phi_n$, we get for all $i \in \{1, \cdots, M\}$,

$$E(1_{\{\tau=t_i\}} \mid \mathcal{G}^{\phi(m)}_{t_i}) \xrightarrow{a.s} 1_{\{\tau=t_i\}}.$$

Define $\tau_m = \min_{\{i \mid E(1_{\{\tau=t_i\}} \mid \mathcal{G}^m_{t_i}) > \frac{1}{2}\}} t_i$. Then

$$\{\tau_m = t_i\} = \{E(1_{\{\tau=t_i\}} \mid \mathcal{G}^m_{t_i}) > \frac{1}{2}\} \cap \{\forall t_j < t_i, E(1_{\{\tau=t_j\}} \mid \mathcal{G}^m_{t_j}) \leq \frac{1}{2}\}$$

and hence τ_m is a $\mathbb{G}^m$ stopping time. Also, obviously, $\tau_{\phi(m)} \xrightarrow{a.s} \tau$, hence $\tau_m \xrightarrow{P} \tau$. ■

1.2.4.3 Weak convergence of σ-fields and the semimartingale property

Assume we are given a sequence of filtrations $(\mathbb{F}^m)_{m\geq 1}$ and define the filtration $\tilde{\mathbb{F}} = (\tilde{\mathcal{F}}_t)_{0\leq t\leq T}$, where $\tilde{\mathcal{F}}_t = \bigvee_m \mathcal{F}^m_t$. We prove in this section a stability result for $\tilde{\mathbb{F}}$ semimartingales. More precisely, we prove that if X is an $\tilde{\mathbb{F}}$ semimartingale, then it remains an $\mathbb{F}$ semimartingale for any limiting (in the sense $\mathcal{F}^m_t \xrightarrow{w} \mathcal{F}_t$, for all $t \in [0, T]$) filtration $\mathbb{F}$ to which it is adapted.

The crucial tool for proving our first theorem is the Bichteler-Dellacherie characterization of semimartingales (see for example [117]). Recall that if $\mathbb{H}$ is a filtration, an $\mathbb{H}$ predictable elementary process H is a process of the form

$$H_t(\omega) = \sum_{i=1}^{k} h_i(\omega) 1_{]t_i,t_{i+1}]}(t);$$

where $0 \le t_1 \le \ldots \le t_{k+1} < \infty$, and each h_i is $\mathcal{H}_{t_i}$ measurable. Moreover, for any $\mathbb{H}$ adapted càdlàg process X and predictable elementary process H of the above form, we write

$$J_X(H) = \sum_{i=1}^{k} h_i(X_{t_{i+1}} - X_{t_i})$$

Theorem 13 (Bichteler-Dellacherie) *Let X be an $\mathbb{H}$ adapted càdlàg process. Suppose that for every sequence $(H_n)_{n\ge1}$ of bounded, $\mathbb{H}$ predictable elementary processes that are null outside a fixed interval $[0, N]$ and convergent to zero uniformly in $(\omega; t)$, we have that $\lim_{n\to\infty} J_X(H_n) = 0$ in probability. Then X is an $\mathbb{H}$ semimartingale.*

The converse is true by the Dominated Convergence Theorem for stochastic integrals. We can now state and prove the main theorem of this subsection.

Theorem 14 *Let $(\mathbb{F}^m)_{m\ge1}$ be a sequence of filtrations. Let $\mathbb{F}$ be a filtration such that for all $t \in [0, T]$, $\mathcal{F}_t^m \xrightarrow{w} \mathcal{F}_t$. Define the filtration $\tilde{\mathbb{F}} = (\tilde{\mathcal{F}}_t)_{0\le t\le T}$, where $\tilde{\mathcal{F}}_t = \bigvee_m \mathcal{F}_t^m$. Let X be an $\mathbb{F}$ adapted càdlàg process such that X is an $\tilde{\mathbb{F}}$ semimartingale. Then X is an $\mathbb{F}$ semimartingale.*

Proof. For a fixed $N > 0$, consider a sequence of bounded, $\mathbb{F}$ predictable elementary processes of the form

$$H_t^n = \sum_{i=1}^{k_n} h_i^n 1_{]t_i^n, t_{i+1}^n]}(t);$$

null outside the fixed time interval $[0, N]$ and with h_i^n being $\mathcal{F}_{t_i^n}$ measurable. Suppose that H^n converges to zero uniformly in (ω, t). We prove that $J_X(H^n) \xrightarrow{P} 0$.

For each m, define the sequence of bounded $\mathbb{F}^m$ predictable elementary processes

$$H_t^{n,m} = \sum_{i=1}^{k_n} E(h_i^n \mid \mathcal{F}_{t_i^n}^m) 1_{]t_i^n, t_{i+1}^n]}(t);$$

By assumption, $\mathcal{F}_t^m \xrightarrow{w} \mathcal{F}_t$ for all $0 \le t \le T$. Hence for all n and $1 \le i \le k_n$, $\mathcal{F}_{t_i^n}^m \xrightarrow{w} \mathcal{F}_{t_i^n}$. Since h_i^n is bounded (hence integrable) and $\mathcal{F}_{t_i^n}$ measurable, it follows from Lemma 12 that $E(h_i^n \mid \mathcal{F}_{t_i^n}^m) \xrightarrow{P} h_i^n$ and hence $E(h_i^n \mid \mathcal{F}_{t_i^n}^m)(X_{t_{i+1}^n} - X_{t_i^n}) \xrightarrow{P} h_i^n(X_{t_{i+1}^n} - X_{t_i^n})$ for each n and $1 \le i \le k_n$ since $(X_{t_{i+1}^n} - X_{t_i^n})$ is finite a.s. Let $\eta > 0$.

$$P\Big(\big|J_X(H^{n,m}) - J_X(H^n)\big| > \eta\Big) \le \sum_{i=1}^{k_n} P\Big(\Big|\big(E(h_i^n \mid \mathcal{F}_{t_i^n}^m) - h_i^n\big)\big(X_{t_{i+1}^n} - X_{t_i^n}\big)\Big| > \frac{\eta}{k_n}\Big)$$

For each fixed n, the right side quantity converges to 0 as m tends to ∞. This proves that for each n,

$$J_X(H^{n,m}) \xrightarrow{P} J_X(H^n).$$

Let $\delta > 0$ and $\varepsilon > 0$. For each n and m,

$$P(|J_X(H^n)| > \delta) \leq P\Big(\big|J_X(H^{n,m}) - J_X(H^n)\big| > \frac{\delta}{2}\Big) + P(|J_X(H^{n,m})| > \frac{\delta}{2}) \tag{1.3}$$

From $J_X(H^{n,m}) \xrightarrow{P} J_X(H^n)$, it follows that for each n, there exists M_0^n such that for all $m \geq M_0^n$,

$$P\Big(\big|J_X(H^{n,m}) - J_X(H^n)\big| > \frac{\delta}{2}\Big) \leq \frac{\varepsilon}{2}$$

Hence $P(|J_X(H^n)| > \delta) \leq \frac{\varepsilon}{2} + P(|J_X(H^{n,M_0^n})| > \frac{\delta}{2})$. First $E(h_i^n \mid \mathcal{F}_{t_i^n}^{M_0^n})$ is bounded, $\tilde{\mathcal{F}}_{t_i^n}$ measurable so that $H_t^{n,M_0^n} = \sum_{i=1}^{k_n} E(h_i^n \mid \mathcal{F}_{t_i^n}^{M_0^n}) 1_{]t_i^n, t_{i+1}^n]}(t)$ is a bounded $\tilde{\mathbb{F}}$ predictable process. Since H^n converges to zero uniformly in (ω, t), it follows that h_i^n converges to zero uniformly in (ω, i) so that there exists n_0 such that for each $n \geq n_0$, for all (ω, i), $|h_i^n(\omega)| \leq \varepsilon$. Hence, for all (ω, t) and $n \geq n_0$

$$|H_t^{n,M_0^n}(\omega)| \leq \sum_{i=1}^{k_n} E(|h_i^n| \mid \mathcal{F}_{t_i^n}^{M_0^n})(\omega) 1_{]t_i^n, t_{i+1}^n]}(t) \leq \varepsilon \sum_{i=1}^{k_n} 1_{]t_i^n, t_{i+1}^n]}(t) \leq \varepsilon$$

Therefore H^{n,M_0^n} is a sequence of bounded $\tilde{\mathbb{F}}$ predictable processes null outside the fixed interval $[0, N]$ that converges uniformly to zero in (ω, t). Since by assumption X is a $\tilde{\mathbb{F}}$ semimartingale, it follows from the converse of Bichteler-Dellacherie's theorem that $J_X(H^{n,M_0^n})$ converges to zero in probability, hence, for n large enough, $P(|J_X(H^{n,M_0^n})| > \frac{\delta}{2}) \leq \frac{\varepsilon}{2}$ and

$$P(|J_X(H^n)| > \delta) \leq \varepsilon$$

Applying now Theorem 13 proves that X is an $\mathbb{F}$ semimartingale. ∎

Let X be an $\tilde{\mathbb{F}}$ semimartingale. Theorem 14 proves that X remains an $\mathbb{F}$ semimartingale for any limiting filtration $\mathbb{F}$ (in the sense $\mathcal{F}_t^m \xrightarrow{w} \mathcal{F}_t$ for all $0 \leq t \leq T$) to which X is adapted. Of course, if $\mathbb{F} \subset \tilde{\mathbb{F}}$, Stricker's theorem (Theorem 5) already implies that X is an $\mathbb{F}$ semimartingale. But there is no general link between the filtration $\tilde{\mathbb{F}} = \bigvee_m \mathbb{F}^m$ and the limiting filtration $\mathbb{F}$. A trivial example is given by taking $\mathbb{F}$ to be the trivial filtration (it can be seen from Definition 4 that the trivial filtration satisfies $\mathcal{F}_t^m \xrightarrow{w} \mathcal{F}_t$, for all t, for any given sequence of filtrations $\mathbb{F}^m$). One can also have $\bigvee_m \mathbb{F}^m \subset \mathbb{F}$, as it is the case in the following important example.

Example 2 *Let X be a càdlàg process. Consider a sequence of subdivisions $\{t_k^n\}$ whose mesh tends to zero and let X^n be the discretized process defined by $X_t^n = X_{t_k^n}$, for all $t_k^n \leq t < t_{k+1}^n$. Let $\mathbb{F}$ and $\mathbb{F}^n$ be the natural filtrations of X and X^n. It is well known that for all t, $\mathcal{F}_{t-} \subset \bigvee_n \mathcal{F}_t^n \subset \mathcal{F}_t$. Also, X^n converges a.s. to the process X, hence $X^n \xrightarrow{P} X$. Assume now that X has no fixed times of discontinuity. Then Lemma 16 guarantees that $\mathcal{F}_t^n \xrightarrow{w} \mathcal{F}_t$, for all t. Moreover, if $\mathbb{F}$ is left-continuous (which is usually the case, and holds for example when X is a càdlàg Hunt Markov process) then $\bigvee_n \mathcal{F}_t^n = \mathcal{F}_t$ for all t.*

We provide now another example where $\bigvee_n \mathcal{F}_t^n$ is itself a limiting σ-field for $(\mathcal{F}_t^n)_{n\geq 1}$, for each t.

Example 3 *Assume that $\mathbb{F}^n$ is a sequence of filtrations such that for all t, the sequence of σ-fields $(\mathcal{F}_t^n)_{n\geq 1}$ is increasing for the inclusion. Define $\tilde{\mathcal{F}}_t = \bigvee_n \mathcal{F}_t^n$. Then for each t, $\mathcal{F}_t^n \stackrel{w}{\to} \tilde{\mathcal{F}}_t$. To see this, fix t and let X be an integrable $\tilde{\mathcal{F}}_t$ measurable random variable. Then $M_n = E(X \mid \mathcal{F}_t^n)$ is a closed martingale and the convergence theorem for closed martingales ensures that M_n converges to X in L^1, which implies that $E(X \mid \mathcal{F}_t^n) \stackrel{P}{\to} X$. Lemma 12 allows us to conclude.*

Checking in practice that X is an $\tilde{\mathbb{F}}$ semimartingale can be a hard task. In sections 1.4 and 1.5, we will replace the strong assumption X *is an* $\tilde{\mathbb{F}}$ *semimartingale* by the more natural assumption X *is an* $\mathbb{F}^n$ *semimartingale, for each* n. Theorem 14 is very instructive since we see from the proof what goes wrong under this new assumption : the change in the order of limits in (1.3) cannot be justified anymore and extra integrability conditions will be needed. These conditions arise naturally in our approach in sections 1.4 and 1.5.

The assumption mentioned above comes out in filtration expansion theory in the following way. Assume we are given a base filtration $\mathbb{F}$ and a sequence of processes N^n which converges (in probability for the Skorohod J_1 topology) to some process N. Let $\mathbb{N}^n$ and $\mathbb{N}$ be their natural filtrations and $\mathbb{G}^n$ (resp. $\mathbb{G}$) the smallest right-continuous filtration containing $\mathbb{F}$ and to which N^n (resp. N) is adapted. Assume that for each n, every $\mathbb{F}$ semimartingale remains a $\mathbb{G}^n$ semimartingale. Does this property also hold between $\mathbb{F}$ and $\mathbb{G}$? We will answer this question in section 1.4 under the assumption of weak convergence of the σ-fields $\mathcal{G}_t^n$ to $\mathcal{G}_t$ for each t, for a class of $\mathbb{F}$ semimartingales X satisfying some integrability conditions. If moreover $\mathbb{G}^n \stackrel{w}{\to} \mathbb{G}$, we are able to provide the $\mathbb{G}$ decomposition of such X.

This ends our mathematical preliminaries section and we start now our filtration expansion study with the progressive filtration expansion with multiple non ranked random times under the assumption that for a given local martingale of the base filtration, one knows what is its decomposition in the initially expanded filtration.

1.3 Progressive filtration expansion under initial-type assumptions

In this section we study progressive filtration expansions with random times. We show how semimartingale decompositions in the expanded filtration can be obtained using a natural link between progressive and initial expansions. The link we exhibit and use is, on an intuitive level, that these two filtrations coincide after the random time. We make this idea precise and use it to establish known and new results in the case of expansion with a single random time. We assume that one knows what the decomposition is in some larger filtration that coincide with the one we are interested in after the random time. Usually, one can take for this larger filtration the initially expanded one. That is the reason we mention in the title *under initial-type assumptions*. The methods are then extended to the multiple time case, without any restrictions on the ordering of the individual times. Finally we study the link between the expanded filtrations from the point of view of filtration shrinkage. As the main analysis progresses, we indicate how the techniques can be generalized to other types of expansions, mainly to what we called in the previous section the (τ, X)-progressive filtration expansion and to progressive filtration expansion with counting processes. These results will be essential for the filtration expansion with a process we focus on in the next section.

1.3.1 Linking progressive and initial filtration expansion with one random time

Assume that a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ is given, and let τ be a random time. Typically τ is not a stopping time with respect to $\mathbb{F}$. We change a bit notations and consider now the larger filtrations $\mathbb{G}^\tau = (\mathcal{G}_t^\tau)_{t\geq 0}$ and $\tilde{\mathbb{G}} = (\tilde{\mathcal{G}}_t)_{t\geq 0}$ given by

$$\mathcal{G}_t^\tau = \bigcap_{u>t} \mathcal{G}_u^{0,\tau} \qquad \text{where} \qquad \mathcal{G}_t^{0,\tau} = \mathcal{F}_t \vee \sigma(\tau \wedge t).$$

and

$$\tilde{\mathcal{G}}_t = \left\{ A \in \mathcal{F}, \exists A_t \in \mathcal{F}_t \mid A \cap \{\tau > t\} = A_t \cap \{\tau > t\} \right\}$$

Now $\mathbb{G}^\tau$ is the *progressive expansion* of $\mathbb{F}$ with τ. Throughout this section $\mathbb{G}$ will denote any right-continuous filtration containing $\mathbb{F}$, making τ a stopping time and satisfying

$$\mathcal{G}_t \cap \{\tau > t\} = \mathcal{F}_t \cap \{\tau > t\} \qquad \text{for all } t \geq 0$$

Our goal in this section is to analyze how $\mathbb{F}$ semimartingales behave in the progressively expanded filtration $\mathbb{G}$. In particular, in case they remain semimartingales in $\mathbb{G}$, we are interested in their canonical decompositions. Under Jacod's assumption (see Assumption 1), this has been done by Jeanblanc and Le Cam [86] in the filtration $\mathbb{G}^\tau$. As one consequence of our approach, we are able to provide a short proof of their main result. Moreover, our technique also works for a larger class of progressively expanded filtrations and under other conditions than Jacod's criterion. This will be illustrated by an example based on the absolute continuity of the Malliavin trace.

The $\mathbb{G}$ decomposition before time τ of an $\mathbb{F}$ local martingale M follows from the classical Jeulin and Yor Theorem as given in Theorem 8. The $\mathbb{G}$ decomposition before time τ follows as a straightforward corollary of Theorem 8 using the following shrinkage result by Föllmer and Protter, see [50].

Lemma 21 *Let $\mathbb{E} \subset \mathbb{F} \subset \mathbb{G}$ be three filtrations. Let X be a $\mathbb{G}$ local martingale. If the optional projection of X onto $\mathbb{E}$ is also a local martingale, then the optional projection of X onto $\mathbb{F}$ is an $\mathbb{F}$ local martingale.*

Putting Theorem 8 and Lemma 21 together provides the $\mathbb{G}$ decomposition before time τ of an $\mathbb{F}$ local martingale M as given in the following theorem.

Theorem 15 *Fix an $\mathbb{F}$ local martingale M. Then*

$$M_{t\wedge\tau} - \int_0^{t\wedge\tau} \frac{d\langle M,\mu\rangle_s + dJ_s}{Z_{s-}}$$

is a $\mathbb{G}$ local martingale. Here, the quantities Z, μ and J are defined as in Theorem 8.

Finding the decomposition after τ is more complicated, but it can be obtained provided that it is known with respect to a suitable auxiliary filtration $\mathbb{H}$. More precisely, one needs that $\mathbb{H}$ and $\mathbb{G}$ coincide after τ, in a certain sense. One such filtration $\mathbb{H}$ when $\mathbb{G}$ is taken to be $\mathbb{G}^\tau$ is, as we will see later, the initial expansion of $\mathbb{F}$ with τ. We now make precise what it means for two filtrations to coincide after τ.

Definition 7 *Let $\mathbb{G}$ and $\mathbb{H}$ be two filtrations such that $\mathbb{G} \subset \mathbb{H}$, and let τ be an $\mathbb{H}$ stopping time. Then $\mathbb{G}$ and $\mathbb{H}$ are said to* coincide after τ *if for every $\mathbb{H}$ optional process X, the process*

$$\mathbf{1}_{[\![\tau,\infty[\![}(X - X_\tau)$$

is $\mathbb{G}$ adapted.

The following lemma establishes some basic properties of filtrations that coincide after τ.

Lemma 22 *Assume that $\mathbb{G}$ and $\mathbb{H}$ are right-continuous and coincide after τ. Then*

(i) For every $\mathbb{H}$ stopping time T, $T \vee \tau$ is $\mathbb{G}$ stopping time. In particular, τ itself is a $\mathbb{G}$ stopping time.

(ii) For every $\mathbb{H}$ optional (predictable) process X, the process $\mathbf{1}_{[\![\tau,\infty[\![}(X - X_\tau)$ is $\mathbb{G}$ optional (predictable).

Proof. For (i), let T be an $\mathbb{H}$ stopping time. Then $T \vee \tau$ is again an $\mathbb{H}$ stopping time, so $X = \mathbf{1}_{[\![0,T\vee\tau]\!]}$ is $\mathbb{H}$ optional. Thus

$$\mathbf{1}_{[\![\tau,\infty[\![}(r)(X_r - X_\tau) = -\mathbf{1}_{\{T\vee\tau<r\}}$$

is $\mathcal{G}_r$-measurable by the assumption that $\mathbb{G}$ and $\mathbb{H}$ coincide after τ. This holds for every $r \geq 0$, so $T \vee \tau$ is a $\mathbb{G}$ stopping time since $\mathbb{G}$ is right-continuous.

For (ii), let X be of the form $X = h\mathbf{1}_{[\![s,t[\![}$ for an $\mathcal{H}_s$-measurable random variable h and fixed $0 \leq s < t$. Then

$$Y_r = \mathbf{1}_{[\![\tau,\infty[\![}(r)(X_r - X_\tau) = \mathbf{1}_{\{\tau\leq r\}}\big(h\mathbf{1}_{\{s\leq r<t\}} - h\mathbf{1}_{\{s\leq\tau<t\}}\big)$$

is $\mathcal{G}_r$-measurable by assumption and defines a càdlàg process. Hence Y is $\mathbb{G}$ optional, and the Monotone Class theorem implies the claim for $\mathbb{H}$ optional processes. The predictable case is similar. ■

Before giving the first result on the $\mathbb{G}$ semimartingale decomposition of an $\mathbb{F}$ local martingale, under the general assumption that such a decomposition is available in some filtration

$\mathbb{H}$ that coincides with $\mathbb{G}$ after τ, we provide examples of such pairs of filtrations $(\mathbb{G}, \mathbb{H})$. For the progressively expanded filtration $\mathbb{G}^\tau$, it turns out that the filtration $\mathbb{H}$ given as the initial expansion of $\mathbb{F}$ with τ coincides with $\mathbb{G}^\tau$ after τ.

Lemma 23 *Let $\mathbb{H}$ be the initial expansion of $\mathbb{F}$ with τ. Then $\mathbb{G}^\tau$ and $\mathbb{H}$ coincide after τ.*

Proof. Let $X = h\mathbf{1}_{[\![s,t[\![}$ for an $\mathcal{H}_s$-measurable random variable h and fixed $0 \leq s < t$. Then

$$\begin{aligned} Y_r = \mathbf{1}_{[\![\tau,\infty[\![}(r)(X_r - X_\tau) &= \mathbf{1}_{\{\tau \leq r\}}\big(h\mathbf{1}_{\{s \leq r < t\}} - h\mathbf{1}_{\{s \leq \tau < t\}}\big) \\ &= h\mathbf{1}_{\{s \leq r < t\}}\mathbf{1}_{\{\tau \leq r\}} - h\mathbf{1}_{\{s \leq \tau < t\}}\mathbf{1}_{\{t \leq r\}} - h\mathbf{1}_{\{s \leq \tau \leq r\}}\mathbf{1}_{\{r < t\}}. \end{aligned}$$

It is enough to prove that the three terms on the right side are $\mathcal{G}^\tau_r$-measurable. Let us consider the first term, the other ones being similar. First let h be of the form $fk(\tau)$ for some $\mathcal{F}_s$-measurable f and Borel function k. Then

$$h\mathbf{1}_{\{s \leq r < t\}}\mathbf{1}_{\{\tau \leq r\}} = fk(\tau)\mathbf{1}_{\{s \leq r < t\}}\mathbf{1}_{\{\tau \leq r\}} = fk(r \wedge \tau)\mathbf{1}_{\{s \leq r < t\}}\mathbf{1}_{\{\tau \leq r\}},$$

which is $\mathcal{G}^\tau_r$-measurable. Using the Monotone Class theorem the result follows for every $\mathcal{H}^0_s$-measurable h, and finally for every $\mathcal{H}_s$-measurable h by a standard argument. ■

As a corollary we may give an alternative characterization of the progressively enlarged filtration $\mathbb{G}^\tau$ as the smallest right-continuous filtration that contains $\mathbb{F}$ and coincides with $\mathbb{H}$ after τ, where $\mathbb{H}$ is the initial expansion of $\mathbb{F}$.

Corollary 3 *Let $\mathbb{H}$ be the initial expansion of $\mathbb{F}$ with τ. Then*

$$\mathbb{G}^\tau = \bigcap \big\{\tilde{\mathbb{G}} : \mathbb{F} \subset \tilde{\mathbb{G}} \subset \mathbb{H},\ \tilde{\mathbb{G}} \textit{ is right-continuous, and } \tilde{\mathbb{G}} \textit{ and } \mathbb{H} \textit{ coincide after } \tau\big\}$$

Proof. To show "$\subset$", let $\tilde{\mathbb{G}}$ be an arbitrary element in the class over which we take the intersection. By Lemma 22, $\tilde{\mathbb{G}}$ is a right-continuous filtration that makes τ a stopping time and contains $\mathbb{F}$. Hence $\mathbb{G}^\tau$ is included in each $\tilde{\mathbb{G}}$ and thus in their intersection. For "$\supset$", Lemma 23 implies that $\mathbb{G}^\tau$ coincides with $\mathbb{H}$ after τ. It is right-continuous, so it is one of the filtrations we are intersecting. ■

Let X be a random variable. Consider now the (τ, X)*-expansion of* $\mathbb{F}$, $\mathbb{G}^{(\tau,X)}$. Recall from Definition 3 that $\mathbb{G}^{(\tau,X)}$ is the smallest right-continuous filtration containing $\mathbb{F}$, that makes τ a stopping time and such that X is $\mathcal{G}^{(\tau,X)}_\tau$ measurable and from Lemma 7 that this filtration is given by

$$\mathcal{G}^{(\tau,X)}_t = \bigcap_{u>t} \mathcal{G}^{0,(\tau,X)}_u \qquad \text{where} \qquad \mathcal{G}^{0,(\tau,X)}_t = \mathcal{F}_t \vee \sigma(\tau \wedge t) \vee \sigma(X1_{\{\tau \leq t\}}).$$

As in Lemma 23 it is easy to prove the following result.

Lemma 24 *Let $\mathbb{H}$ be the initial expansion of $\mathbb{F}$ with (τ, X). Then $\mathbb{G}^{(\tau,X)}$ coincides with $\mathbb{H}$ after τ.*

Also, $\mathbb{G}^{(\tau,X)}$ satisfies the crucial condition $\mathcal{G}_t^{(\tau,X)} \cap \{\tau > t\} = \mathcal{F}_t \cap \{\tau > t\}$ for all $t \geq 0$ as proved in Lemma 8.

We are now ready to give the first result on the $\mathbb{G}$ semimartingale decomposition of an $\mathbb{F}$ local martingale, under the general assumption that such a decomposition is available in some filtration $\mathbb{H}$ that coincides with $\mathbb{G}$ after τ.

Theorem 16 *Let M be an $\mathbb{F}$ local martingale. Let $\mathbb{G}$ be any progressive expansion of $\mathbb{F}$ with τ satisfying $\mathcal{G}_t \cap \{\tau > t\} = \mathcal{F}_t \cap \{\tau > t\}$ for all $t \geq 0$ and let $\mathbb{H}$ be a filtration that coincides with $\mathbb{G}$ after τ. Suppose there exists an $\mathbb{H}$ predictable finite variation process A such that $M - A$ is an $\mathbb{H}$ local martingale. Then M is a $\mathbb{G}$ semimartingale, and*

$$M_t - \int_0^{t\wedge\tau} \frac{d\langle M, \mu\rangle_s + dJ_s}{Z_{s-}} - \int_{t\wedge\tau}^t dA_s$$

is the local martingale part of its $\mathbb{G}$ decomposition. Here Z, μ and J are defined as in Theorem 8.

Proof. The process $M_t^{\mathbb{G}} = M_{t\wedge\tau} - \int_0^{t\wedge\tau} \frac{d\langle M,\mu\rangle_s + dJ_s}{Z_{s-}}$ is a $\mathbb{G}$ local martingale by the Jeulin-Yor theorem (Theorem 15). Next, define

$$M^{\mathbb{H}} = \mathbf{1}_{[\![\tau,\infty[\![}(\tilde{M} - \tilde{M}_\tau),$$

where $\tilde{M}_t = M_t - A_t$. Since $\tilde{M}$ is an $\mathbb{H}$ local martingale, $M^{\mathbb{H}}$ is also. Moreover, if $(T_n)_{n\geq 1}$ is a sequence of $\mathbb{H}$ stopping times that reduce $\tilde{M}$, then $T_n' = T_n \vee \tau$ yields a reducing sequence for $M^{\mathbb{H}}$. Lemma 22 (i) shows that the T_n' are in fact $\mathbb{G}$ stopping times, and since $\mathbb{G}$ and $\mathbb{H}$ coincide after τ, $M^{\mathbb{H}}$ is $\mathbb{G}$ adapted. This implies that $M^{\mathbb{H}}_{\cdot\wedge T_n'}$ is an $\mathbb{H}$ martingale that is $\mathbb{G}$ adapted, and is therefore a $\mathbb{G}$ martingale. It follows that $M^{\mathbb{H}}$ is a $\mathbb{G}$ local martingale. It now only remains to observe that

$$M_t^{\mathbb{G}} + M_t^{\mathbb{H}} = M_t - \int_0^{t\wedge\tau} \frac{d\langle M, \mu\rangle_s + dJ_s}{Z_{s-}} - \int_{t\wedge\tau}^t dA_s,$$

which thus is a $\mathbb{G}$ local martingale. Finally, by Lemma 22 (ii), the last term is $\mathbb{G}$ predictable, so we obtain indeed the $\mathbb{G}$ semimartingale decomposition. ■

Part of the proof of Theorem 16 can be viewed as a statement about filtration shrinkage. According to a result by Föllmer and Protter [50], if $\mathbb{G} \subset \mathbb{H}$ are two nested filtrations and L is an $\mathbb{H}$ local martingale that can be reduced using $\mathbb{G}$ stopping times, then its optional projection onto $\mathbb{G}$ is again a local martingale. In our case L corresponds to $M^{\mathbb{H}}$, which is $\mathbb{G}$ adapted and hence coincides with its optional projection.

We now proceed to examine two particular situations where $\mathbb{G}$ and $\mathbb{H}$ coincide after τ, and where the $\mathbb{H}$ decomposition $M - A$ is available. First we make an absolute continuity assumption on the $\mathcal{F}_t$ conditional laws of τ or (τ, X), as in Assumption 1. We then assume that $\mathbb{F}$ is a Wiener filtration, and impose a condition related to the Malliavin derivatives of the process of $\mathcal{F}_t$ conditional distributions. This is based on theory developed by Imkeller, Pontier and Weisz [71] and Imkeller [70].

1.3.1.1 Jacod's criterion

In this section we study the case where τ or (τ, X) satisfy *Jacod's criterion*, which was recalled in section 1.2.

Theorem 17 *Let M be an $\mathbb{F}$ local martingale.*

(i) If τ satisfies Jacod's Criterion (Assumption 1), then M is a $\mathbb{G}^\tau$ semimartingale.

(ii) Let X be a random variable such that (τ, X) satisfies Jacod's Criterion (Assumption 1), then M is a $\mathbb{G}^{(\tau,X)}$ semimartingale.

Proof. We prove (i). Let $\mathbb{H}^\tau$ be the initial expansion of F with τ. It follows from Jacod's theorem (see Theorems VI.10 and VI.11 in [7]) that M is an $\mathbb{H}^\tau$ semimartingale, which is $\mathbb{G}^\tau$ adapted. It is also a $\mathbb{G}^\tau$ semimartingale by Theorem 5. The proof of (ii) is similar. ■

We provide the explicit decompositions using the classical result by Jacod recalled in Theorem 6 in the mathematical preliminaries of this chapter. Immediate consequences of Theorem 16 and Theorem 6 are the following corollaries.

Corollary 4 *Let M be an $\mathbb{F}$ local martingale, and assume that τ satisfies Assumption 1. Then M is a $\mathbb{G}^\tau$ semimartingale, and*

$$M_t - \int_0^{t\wedge\tau} \frac{d\langle M, \mu\rangle_s + dJ_s}{Z_{s-}} - \int_{t\wedge\tau}^t k_s(\tau) dA_s$$

is the local martingale part of its $\mathbb{G}^\tau$ decomposition. Here Z, μ and J are defined as in Theorem 8, and k and A as in Theorem 6 with $d = 1$ and $\xi = \tau$.

Notice that this recovers the main result in [86] (Theorem 3.1), since by Theorem 6 (*ii*) we may write

$$\int_{t\wedge\tau}^t k_s(\tau) dA_s = \int_{t\wedge\tau}^t \frac{d\langle p(u), M\rangle_s}{p_{s-}(u)}\bigg|_{u=\tau},$$

whenever the right side makes sense. See [86] for a detailed discussion.

Corollary 5 *Let M be an $\mathbb{F}$ local martingale. Let X be a random variable and assume that (τ, X) satisfies Assumption 1. Then M is a $\mathbb{G}^{(\tau,X)}$ semimartingale, and*

$$M_t - \int_0^{t\wedge\tau} \frac{d\langle M, \mu\rangle_s + dJ_s}{Z_{s-}} - \int_{t\wedge\tau}^t k_s(\tau, X) dA_s$$

is the local martingale part of its $\mathbb{G}^{(\tau,X)}$ decomposition. Here Z, μ and J are defined as in Theorem 8, and k and A as in Theorem 6 with $d = 2$ and $\xi = (\tau, X)$.

1.3.1.2 Absolute continuity of the Malliavin trace

In two papers on models for insider trading in mathematical finance, Imkeller et al. [71] and Imkeller [70] introduced an extension of Jacod's criterion for initial expansions, based on

Malliavin calculus. Given a measure-valued random variable $F(du;\omega)$ defined on Wiener space with coordinate process $(W_t)_{0\le t\le 1}$, they introduce a Malliavin derivative $D_tF(du;\omega)$, defined for all F satisfying certain regularity conditions. The full details are outside the scope of the present subsection, and we refer the interested reader to [71] and [70]. We continue to let $\mathbb{H}$ be the initial expansion of $\mathbb{F}$.

The extension of Jacod's criterion is the following. Let $P_t(du,\omega) = P(\tau \in du \mid \mathcal{F}_t)(\omega)$, and assume that $D_tP_t(du,\omega)$ exists and satisfies

$$\sup_{f\in C_b(\mathbb{R}),\ \|f\|\le 1} E\left(\int_0^1 \langle D_sP_s(du), f\rangle^2 ds\right) < \infty.$$

Here $C_b(\mathbb{R})$ is the space of bounded and continuous functions on $\mathbb{R}$, $\|\cdot\|$ is the supremum norm, and $\langle F(du), f\rangle = \int_{\mathbb{R}_+} f(u)F(du)$ for a random measure F and $f \in C_b(\mathbb{R})$. Assume also that

$$D_tP_t(du,\omega) \ll P_t(du,\omega) \quad a.s. \text{ for all } t \in [0,1],$$

and let $g_t(u;\omega)$ be a suitably measurable version of the corresponding density. Then they prove the following result.

Theorem 18 *Under the above conditions, if $\int_0^1 |g_t(\tau)|dt < \infty$ a.s., then*

$$W_t - \int_0^t g_s(\tau)ds$$

is a Brownian motion in the initially expanded filtration $\mathbb{H}$.

One example where this holds but Jacod's criterion fails is $\tau = \sup_{0\le t\le 1} W_t$. In this case $g_t(\tau)$ can be computed explicitly and the $\mathbb{H}$ decomposition of W obtained. Due to the martingale representation theorem in $\mathbb{F}$, this allows one to obtain the $\mathbb{H}$ decomposition for every $\mathbb{F}$ local martingale. Using Theorem 16, the decomposition in the progressively expanded filtration $\mathbb{G}$ can then also be obtained.

Corollary 6 *Under the assumptions of Theorem 18,*

$$W_t - \int_0^{t\wedge\tau} \frac{d\langle W, Z\rangle_s}{Z_{s-}} - \int_{t\wedge\tau}^t g_s(\tau)ds$$

is a $\mathbb{G}$ Brownian motion.

1.3.2 The case of multiple non ranked random times

We now move on to progressive expansions with multiple random times. We start again with a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$, but instead of a single random time we consider a vector of random times

$$\boldsymbol{\tau} = (\tau_1, \ldots, \tau_n).$$

We emphasize that there are no restrictions on the ordering of the individual times. This is a significant departure from previous work in the field, where the times are customarily assumed to be ordered. The progressive expansion of $\mathbb{F}$ with $\boldsymbol{\tau}$ is

$$\mathcal{G}_t = \bigcap_{u>t} \mathcal{G}_u^0 \qquad \text{where} \qquad \mathcal{G}_t^0 = \mathcal{F}_t \vee \sigma(\tau_i \wedge t; i = 1, \ldots, n),$$

and we are interested in the semimartingale decompositions of $\mathbb{F}$ local martingales in the $\mathbb{G}$ filtration. Several other filtrations will also appear, and we now introduce notation that will be in place for the remainder of this section, except in Theorem 20 and its corollary. Let $I \subset \{1, \ldots, n\}$ be an index set.

- $\sigma_I = \max_{i \in I} \tau_i$ and $\rho_I = \min_{j \notin I} \tau_j$
- $\mathbb{G}^I$ denotes the initial expansion of $\mathbb{F}$ with the random vector $\boldsymbol{\tau}_I = (\tau_i)_{i \in I}$
- $\mathbb{H}^I$ denotes the progressive expansion of $\mathbb{G}^I$ with the random time ρ_I

If $I = \emptyset$, then $\mathbb{G}^I = \mathbb{F}$ and $\mathbb{H}^I$ is the progressive expansion of $\mathbb{F}$ with $\rho_\emptyset = \min_{i=1,\ldots,n} \tau_i$. If on the other hand $I = \{1, \ldots, n\}$, then $\mathbb{G}^I = \mathbb{H}^I$, and coincides with the initial expansion of $\mathbb{F}$ with $\boldsymbol{\tau}_I = \boldsymbol{\tau}$. Notice also that we always have $\mathbb{G}^I \subset \mathbb{H}^I$.

The idea from Section 1.3.1 can be modified to work in the present context. The intuition is that the filtrations $\mathbb{G}$ and $\mathbb{H}^I$ coincide on $[\![\sigma_I, \rho_I[\![$. The $\mathbb{G}$ decomposition on $[\![\sigma_I, \rho_I[\![$ of an $\mathbb{F}$ local martingale M can then be obtained by computing its decomposition in $\mathbb{H}^I$. This is done in two steps. First it is obtained in $\mathbb{G}^I$ using, for instance, Jacod's theorem (Theorem 6), and then in $\mathbb{H}^I$ up to time ρ_I using the Jeulin-Yor theorem (Theorem 8).

The following results collect some properties of the relationship between $\mathbb{H}^I$ and $\mathbb{G}$, thereby clarifying in which sense they coincide on $[\![\sigma_I, \rho_I[\![$. We take the index set I to be given and fixed.

Lemma 25 *Let X be an $\mathcal{H}_t^I$-measurable random variable. Then the quantity $X\mathbf{1}_{\{\sigma_I \le t\}}$ is $\mathcal{G}_t$-measurable. As a consequence, if H is an $\mathbb{H}^I$ optional (predictable) process, then $\mathbf{1}_{[\![\sigma_I, \rho_I[\![}(H - H_{\sigma_I})$ is $\mathbb{G}$ optional (predictable).*

Proof. Let X be of the form $X = fk(\rho_I \wedge t) \prod_{i \in I} h_i(\tau_i)$ for some $\mathcal{F}_t$-measurable random variable f and Borel functions k, h_i. Then

$$X\mathbf{1}_{\{\sigma_I \le t\}} = fk(\rho_I \wedge t) \prod_{i \in I} h_i(\tau_i \wedge t)\mathbf{1}_{\{\sigma_I \le t\}},$$

which is $\mathcal{G}_t$-measurable, since σ_I and ρ_I are $\mathbb{G}$ stopping times and $\tau_i \wedge t$ is $\mathcal{G}_t$-measurable by construction. The Monotone Class theorem shows that the statement holds for every X that is measurable for $\mathcal{F}_t \vee \sigma(\tau_i : i \in I) \vee \sigma(\rho_I \wedge t)$. $\mathbb{H}^I$ is the right-continuous version of this filtration, so the result follows.

Now consider an $\mathbb{H}^I$ predictable process of the form $H = h\mathbf{1}_{]\!]s,t]\!]}$ with $s \le t$ and h an $\mathcal{H}_s^I$-measurable random variable. Then

$$H_r\mathbf{1}_{[\![\sigma_I, \rho_I[\![}(r) = h\mathbf{1}_{\{s<r\le t\}}\mathbf{1}_{\{\sigma_I \le r < \rho_I\}} - h\mathbf{1}_{\{s<\sigma_I\le r<\rho_I\}},$$

which defines a $\mathbb{G}$ predictable process using the first part of the lemma. An application of the Monotone Class theorem yields the desired result in the predictable case. The optional case is similar. ■

Lemma 26 *Let T_n be an $\mathbb{H}^I$ stopping time and define $T'_n = (\sigma_I \vee T_n) \wedge (\rho_I \vee n)$. Then T'_n is a $\mathbb{G}$ stopping time.*

Proof. We need to show that $\{T'_n > t\} \in \mathcal{G}_t$ for every $t \geq 0$. Careful set manipulations yield

$$\begin{aligned}\mathbf{1}_{\{T'_n > t\}} &= \mathbf{1}_{\{\sigma_I \vee T_n > t\}} \mathbf{1}_{\{\rho_I \vee n > t\}} \\ &= \left(1 - \mathbf{1}_{\{\sigma_I \leq t\}} \mathbf{1}_{\{T_n \leq t\}}\right) \left(1 - \mathbf{1}_{\{\rho_I \leq t\}} \mathbf{1}_{\{n \leq t\}}\right).\end{aligned}$$

We have that $1 - \mathbf{1}_{\{\rho_I \leq t\}} \mathbf{1}_{\{n \leq t\}} = 1 - \mathbf{1}_{\{\rho_I \leq t\}}(1 - \mathbf{1}_{\{n > t\}}) = \mathbf{1}_{\{\rho_I > t\}} + \mathbf{1}_{\{\rho_I \leq t\}} \mathbf{1}_{\{n > t\}}$, so

$$\begin{aligned}\mathbf{1}_{\{T'_n > t\}} = \mathbf{1}_{\{\rho_I > t\}} &- \mathbf{1}_{\{\sigma_I \leq t\}} \mathbf{1}_{\{T_n \leq t\}} \mathbf{1}_{\{\rho_I > t\}} \\ &+ \mathbf{1}_{\{\rho_I \leq t\}} \mathbf{1}_{\{n > t\}} - \mathbf{1}_{\{\sigma_I \leq t\}} \mathbf{1}_{\{T_n \leq t\}} \mathbf{1}_{\{\rho_I \leq t\}} \mathbf{1}_{\{n > t\}}.\end{aligned}$$

The first and third terms are $\mathcal{G}_t$-measurable since ρ_I is a $\mathbb{G}$ stopping time. The second term is equal to $\mathbf{1}_{\{T_n \leq t\}} \mathbf{1}_{\{\sigma_I \leq t < \rho_I\}}$ and is thus $\mathcal{G}_t$-measurable by Lemma 25, which also gives the measurability of the fourth term. ■

The next theorem generalizes Theorem 16 to the case of progressive enlargement with multiple, not necessarily ranked times.

Theorem 19 *Let M be an $\mathbb{F}$ local martingale such that $M_0 = 0$. For each $I \subset \{1, \ldots, n\}$, suppose there exists a $\mathbb{G}^I$ predictable finite variation process A^I such that $M - A^I$ is a $\mathbb{G}^I$ local martingale. Moreover, define*

$$Z^I_t = P(\rho_I > t \mid \mathcal{G}^I_t),$$

and let μ^I and J^I be as in Theorem 8. Then M is a $\mathbb{G}$ semimartingale with local martingale part $M_t - A_t$, where

$$A_t = \sum_I \mathbf{1}_{\{\sigma_I \leq \rho_I\}} \left(\int_{t \wedge \sigma_I}^{t \wedge \rho_I} dA^I_s + \int_{t \wedge \sigma_I}^{t \wedge \rho_I} \frac{d\langle M, \mu^I \rangle_s + dJ^I_s}{Z^I_{s-}} \right).$$

Here the sum is taken over all $I \subset \{1, \ldots, n\}$.

Notice that when $I = \emptyset$, then $\sigma_I = 0$ and $\mathbb{G}^I = \mathbb{F}$, so that $A^I = 0$ and does not contribute to A_t. Similarly, when $I = \{1, \ldots, n\}$, $\rho_I = \infty$ and we have $Z^I_t = 1$, causing both $\langle M, \mu^I \rangle$ and J to vanish. Note also that it is also a classical fact that $Z^I_{t-} > 0$ on $\{\rho_I \geq t\}$.

Before proving Theorem 19 we need the two following technical lemmas.

Lemma 27 *For any process M, the following identity holds.*

$$\sum_{I \subset \{1, \ldots, n\}} \mathbf{1}_{\{\sigma_I \leq \rho_I\}} (M_{t \wedge \rho_I} - M_{t \wedge \sigma_I}) = M_t - M_0$$

Proof. Fix $\omega \in \Omega$ and let $0 \leq \tau_{\eta(1)} \leq \cdots \tau_{\eta(n)}$ be the ordered times. Then for I of the form $I = \{\eta(1), \cdots, \eta(k)\}$ (for some $0 \leq k \leq n$, with the convention $I = \emptyset$ when $k = 0$), one has

$$1_{\{\sigma_I \leq \rho\}}(M_{t\wedge\rho_I} - M_{t\wedge\sigma_I}) = M_{t\wedge\tau_{\eta(k+1)}} - M_{t\wedge\tau_{\eta(k)}}$$

with the convention that $\tau_{n+1} = \infty$. For other I the term is zero, since in that case $\sigma_I \geq \rho_I$. Summing over all subsets of $\{1, \cdots, n\}$ provides the result.

■

Lemma 28 *Let L be a local martingale in some filtration $\mathbb{F}$, suppose σ and ρ are two stopping times, and define a process $N_t = \mathbf{1}_{\{\sigma\leq\rho\}}(L_{t\wedge\rho} - L_{t\wedge\sigma})$.*

(i) N is again a local martingale.

(ii) Let T be a stopping time and define $T' = (\sigma \vee T) \wedge (\rho \vee n)$ for a fixed n. Then $N_{t\wedge T} = N_{t\wedge T'}$.

Proof. *Part* (*i*)*:* This is clear since N can be written

$$N_t = \mathbf{1}_{\{\sigma\leq\rho\}}(L_{t\wedge\rho} - L_{t\wedge\sigma}) = L_{t\wedge\rho} - L_{t\wedge\sigma\wedge\rho}$$

Part (*ii*)*:* The proof consists of a careful analysis of the interplay between the indicators involved in the definition of N and T'. On $\{\sigma \leq \rho\}$ we have

$$t \wedge T' \wedge \rho = t \wedge (\sigma \vee T) \wedge (\rho \vee n) \wedge \rho = t \wedge (\sigma \vee T) \wedge \rho$$

and

$$t \wedge T' \wedge \sigma = t \wedge (\sigma \vee T) \wedge (\rho \vee n) \wedge \sigma = t \wedge (\sigma \vee T) \wedge \sigma.$$

On $\{T \leq \sigma\} \cap \{\sigma \leq \rho\}$, these are both equal to $t \wedge \sigma$, so on this set,

$$N_{t\wedge T'}\mathbf{1}_{\{T\leq\sigma\}} = \mathbf{1}_{\{\sigma\leq\rho\}}\mathbf{1}_{\{T\leq\sigma\}}(L_{t\wedge\sigma} - L_{t\wedge\sigma}) = 0.$$

Hence $N_{t\wedge T'} = N_{t\wedge T'}\mathbf{1}_{\{T>\sigma\}}$. Similarly, $N_{t\wedge T} = 0$ on $\{T \leq \sigma\}$, and hence $N_{t\wedge T}\mathbf{1}_{\{T>\sigma\}} = N_{t\wedge T}$. But on $\{T > \sigma\} \cap \{\sigma \leq \rho\}$,

$$t \wedge (\sigma \vee T') \wedge \rho = t \wedge T \wedge \rho \qquad \text{and} \qquad t \wedge (\sigma \vee T') \wedge \sigma = t \wedge T \wedge \sigma,$$

so $N_{t\wedge T'}\mathbf{1}_{\{T>\sigma\}} = N_{t\wedge T}\mathbf{1}_{\{T>\sigma\}}$. This yields $N_{t\wedge T'} = N_{t\wedge T}$ as required. ■

Proof of Theorem 19. For each fixed index set I, the process $M_t - \int_0^t dA_s^I$ is a local martingale in the initially expanded filtration $\mathbb{G}^I$ by assumption. Now, $\mathbb{H}^I$ is obtained from $\mathbb{G}^I$ by a progressive expansion with ρ_I, so Theorem 8 yields that

$$M_t^I = M_{t\wedge\rho_I} - \int_0^{t\wedge\rho_I} dA_s^I - \int_0^{t\wedge\rho_I} \frac{d\langle M, \mu^I\rangle_s + dJ_s^I}{Z_{s-}^I}$$

is an $\mathbb{H}^I$ local martingale. Define the process

$$N_t^I = \mathbf{1}_{\{\sigma_I\leq\rho_I\}}\left(M_{t\wedge\rho_I}^I - M_{t\wedge\sigma_I}^I\right).$$

Our goal is to prove that N^I is a $\mathbb{G}$ local martingale. This will imply the statement of the theorem, since summing the N^I over all index sets I and using Lemma 27 yields precisely $M - A$.

By part (i) of Lemma 28, N^I is a local martingale in $\mathbb{H}^I$. Write

$$N_t^I = \mathbf{1}_{\{\sigma_I \leq t < \rho_I\}} \left(M_t^I - M_{\sigma_I}^I\right) + \mathbf{1}_{\{\sigma_I \leq \rho_I < t\}} \left(M_{\rho_I}^I - M_{\sigma_I}^I\right) = Y_1 + Y_2$$

and apply Lemma 25 with $X = Y_1$ for the first term and $X = Y_2$ for the second to see that N^I is in fact $\mathbb{G}$ adapted. The use of Lemma 25 is valid because both Y_1 and Y_2 are $\mathcal{H}_t^I$-measurable.

Next, let $(T_n)_{n\geq 1}$ be a reducing sequence for N^I in $\mathbb{H}^I$. By Lemma 26 we know that $T_n' = (\sigma_I \vee T_n) \wedge (\rho_I \vee n)$ are $\mathbb{G}$ stopping times, and since $T_n' \geq T_n \wedge n$ we have $T_n' \uparrow \infty$ a.s. Moreover, part (ii) of Lemma 28 implies that $N_{t\wedge T_n}^I = N_{t\wedge T_n'}^I$. Hence $(N_{t\wedge T_n'}^I)_{t\geq 0}$ is an $\mathbb{H}^I$ martingale that is $\mathbb{G}$ adapted, and therefore even a $\mathbb{G}$ martingale. We deduce from the above that N^I is a $\mathbb{G}$ local martingale. A final application of Lemma 25 shows that A is $\mathbb{G}$ predictable, so we obtain indeed the $\mathbb{G}$ semimartingale decomposition of M. This completes the proof of Theorem 19. ∎

We now proceed to study the special case where the vector $\boldsymbol{\tau}$ of random times satisfies Jacod's criterion, meaning that Assumption 1 holds, now with state space $E = \mathbb{R}_+^n$. Again there is no loss of generality to let η be the law of $\boldsymbol{\tau}$. We will further assume that η is absolutely continuous w.r.t to Lebesgue measure, so that $\eta(d\boldsymbol{u}) = h(\boldsymbol{u})d\boldsymbol{u}$. Provided the law of $\boldsymbol{\tau}$ does not have atoms, this restriction is not essential—everything that follows goes through without it—but it simplifies the already quite cumbersome notation.

The joint $\mathcal{F}_t$ conditional density corresponding to this choice of η is denoted $p_t(\boldsymbol{u};\omega)$. That is,

$$P(\boldsymbol{\tau} \in d\boldsymbol{u} \mid \mathcal{F}_t) = p_t(\boldsymbol{u})du_1 \cdots du_n,$$

where we suppressed the dependence on ω. Now, for an index set $I \subset \{1, \ldots, n\}$ with $|I| = m$, we have, for $\boldsymbol{u}_I \in \mathbb{R}_+^m$,

$$P(\boldsymbol{\tau}_I \leq \boldsymbol{u}_I \mid \mathcal{F}_t) = \int_{\boldsymbol{v}_I \leq \boldsymbol{u}_I} \int_{\boldsymbol{v}_{-I} \geq 0} p_t(\boldsymbol{v}_I; \boldsymbol{v}_{-I}) d\boldsymbol{v}_{-I} d\boldsymbol{v}_I,$$

where $\boldsymbol{\tau}_I$ is the subvector of $\boldsymbol{\tau}$ whose components have indices in I, and where $\boldsymbol{v}_I$ and $\boldsymbol{v}_{-I}$ are the subvectors of $\boldsymbol{v}$ with indices in I, respectively not in I. Inequalities should be interpreted componentwise. The above shows that $\boldsymbol{\tau}_I$ also satisfies Assumption 1, so that there is an appropriately measurable function $p_t^I(\boldsymbol{u}_I;\omega)$ such that

$$P(\boldsymbol{\tau}_I \in d\boldsymbol{u}_I \mid \mathcal{F}_t) = p_t^I(\boldsymbol{u}_I)d\boldsymbol{u}_I.$$

Moreover, this conditional density p^I is given by

$$p_t^I(\boldsymbol{u}_I) = \int_{\mathbb{R}_+^{n-m}} p_t(\boldsymbol{u}_I; \boldsymbol{u}_{-I}) d\boldsymbol{u}_{-I}.$$

Define

$$p_t(\boldsymbol{u}_{-I} \mid \boldsymbol{\tau}_I) = \frac{p_t(\boldsymbol{\tau}_I; \boldsymbol{u}_{-I})}{\int_0^\infty \cdots \int_0^\infty p_t(\boldsymbol{\tau}_I; \boldsymbol{v}_{-I}) d\boldsymbol{v}_{-I}}.$$

This is the conditional density of $\boldsymbol{\tau}_{-I}$ given $\mathcal{F}_t \vee \sigma(\tau_I)$, as one can verify using standard arguments. The quantity above is well defined, since the denominator only vanishes if $p_t(\boldsymbol{\tau}_I; \boldsymbol{u}_{-I})$ is identically zero for each $\boldsymbol{u}_{-I}$, in which case the numerator and the conditional density are also null. Defining

$$Z_t^I = \int_t^\infty \cdots \int_t^\infty p_t(\boldsymbol{u}_{-I} \mid \boldsymbol{\tau}_I) d\boldsymbol{u}_{-I}, \tag{1.4}$$

we have $Z_t^I = P(\rho_I > t \mid \mathcal{F}_t \vee \sigma(\tau_I))$. One then readily checks that we also have

$$Z_t^I = P(\rho_I > t \mid \mathcal{G}_t^I).$$

We can now state the decomposition theorem for continuous $\mathbb{F}$ local martingales in the progressively expanded filtration $\mathbb{G}$, when Jacod's criterion is satisfied. Recall that $\sigma_I = \max_{i \in I} \tau_i$ and $\rho_I = \min_{j \notin I} \tau_j$.

Corollary 7 *Let M be an $\mathbb{F}$ local martingale, and assume Assumption 1 is satisfied for $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_n)$. Then M is a $\mathbb{G}$ semimartingale. For each $I \subset \{1, \ldots, n\}$, let Z^I be given by (1.4) and let μ^I and J^I be as in Theorem 8. Then $M_t - A_t$ is a $\mathbb{G}$ local martingale, where*

$$A_t = \sum_I \mathbf{1}_{\{\sigma_I \le \rho_I\}} \left(\int_{t \wedge \sigma_I}^{t \wedge \rho_I} \frac{d\langle p^I(\boldsymbol{u}_I), M\rangle_s}{p_{s-}^I(\boldsymbol{u}_I)}\Big|_{\boldsymbol{u}_I = \boldsymbol{\tau}_I} + \int_{t \wedge \sigma_I}^{t \wedge \rho_I} \frac{d\langle M, \mu^I\rangle_s + dJ_s^I}{Z_{s-}^I} \right).$$

Here the sum is taken over all $I \subset \{1, \ldots, n\}$.

Proof. We apply Theorem 19 and notice that it follows from Theorem 6 that

$$A_t^I = \int_0^t k_s^I(\boldsymbol{\tau}_I) d\langle M, M\rangle_s = \int_0^t \frac{d\langle p^I(\boldsymbol{u}_I), M\rangle_s}{p_{s-}^I(\boldsymbol{u}_I)}\Big|_{\boldsymbol{u}_I = \boldsymbol{\tau}_I}.$$

This completes the proof. ■

We end this section by pointing out that the filtration $\mathbb{G}^{(\tau, X)}$ introduced earlier can be generalized to the multiple time case. To state the precise result, we first suppose that each random time τ_i is accompanied by a random variable X_i, and let $\boldsymbol{X} = (X_1, \ldots, X_n)$. Define the filtration $\mathbb{G}^{(\boldsymbol{\tau}, \boldsymbol{X})}$ by

$$\mathcal{G}_t^{(\boldsymbol{\tau}, \boldsymbol{X})} = \bigcap_{u > t} \mathcal{G}_u^{0, (\boldsymbol{\tau}, \boldsymbol{X})},$$

where

$$\mathcal{G}_t^{0, (\boldsymbol{\tau}, \boldsymbol{X})} = \mathcal{F}_t \vee \sigma(\tau_i \wedge t : i = 1, \ldots, n) \vee \sigma(X_i 1_{\{\tau_i \le t\}} : i = 1, \ldots, n).$$

Let $I \subset \{1, \ldots, n\}$ be an index set. Assume for simplicity that $P(\tau_i = \tau_j) = 0$ for $i \neq j$. We may then define

- $\boldsymbol{X}_I = (X_i)_{i\in I}$.
- $Y_I = X_{i^*}$, where $i^* \in I$ is the index for which $\rho_I = \tau_{i^*}$.

For the statement and proof of Theorem 20, we redefine the objects $\mathbb{G}^I$ and $\mathbb{H}^I$ as follows. For an index set $I \subset \{1, \ldots, n\}$,

- $\mathbb{G}^I$ denotes the initial expansion of $\mathbb{F}$ with the random vector $(\boldsymbol{\tau}_I, \boldsymbol{X}_I) = (\tau_i, X_i)_{i\in I}$.
- $\mathbb{H}^I$ denotes the (ρ_I, Y_I)-expansion of $\mathbb{G}^I$.

Theorem 20 *Let M be an $\mathbb{F}$ local martingale such that $M_0 = 0$. For each $I \subset \{1, \ldots, n\}$, suppose there exists a $\mathbb{G}^I$ predictable finite variation process A^I such that $M - A^I$ is a $\mathbb{G}^I$ local martingale. Moreover, define*

$$Z_t^I = P(\rho_I > t \mid \mathcal{G}_t^I),$$

and let μ^I and J^I be as in Theorem 8. Then M is a $\mathbb{G}^{(\boldsymbol{\tau},\boldsymbol{X})}$ semimartingale with local martingale part $M_t - A_t$, where

$$A_t = \sum_I \mathbf{1}_{\{\sigma_I \le \rho_I\}} \left(\int_{t\wedge\sigma_I}^{t\wedge\rho_I} dA_s^I + \int_{t\wedge\sigma_I}^{t\wedge\rho_I} \frac{d\langle M, \mu^I\rangle_s + dJ_s^I}{Z_{s-}^I} \right).$$

Here the sum is taken over all $I \subset \{1, \ldots, n\}$.

Proof. The proof is the same as that of Theorem 19, except for the following points: Instead of Theorem 8, we invoke Theorem 15, which is justified by Lemma 8. Moreover, it must be verified that Lemma 25 remains valid for the redefined $\mathbb{H}^I$ and $\mathbb{G} = \mathbb{G}^{(\boldsymbol{\tau},\boldsymbol{X})}$. This is easily done: in the proof of Lemma 25, simply replace $X = fk(\rho_I \wedge t)\prod_{i\in I} h_i(\tau_i)$ by

$$X = fk(\rho_I \wedge t)\ell(Y_I \mathbf{1}_{\{\rho_I \le t\}}) \prod_{i\in I} h_i(\tau_i) g_i(X_i \mathbf{1}_{\{\tau_i \le t\}}),$$

where $\ell(\cdot)$ and $g_i(\cdot)$ are Borel functions. ■

Define the process

$$N_t^n = \sum_{i=1}^n X_i 1_{\{\tau_i \le t\}}$$

Let $\mathbb{N}^n$ be the smallest right continuous filtration containing $\mathbb{F}$ and to which the process N^n is adapted. Under the same assumption as in Theorem 20, $\mathbb{F}$ semimartingales remain $\mathbb{N}^n$ semimartingales.

Corollary 8 *Let M be an $\mathbb{F}$ local martingale such that $M_0 = 0$. For each $I \subset \{1, \ldots, n\}$, suppose there exists a $\mathbb{G}^I$ predictable finite variation process A^I such that $M - A^I$ is a $\mathbb{G}^I$ local martingale. Then M is a $\mathbb{N}^n$ semimartingale.*

Proof. Applying Theorem 20, M is a $\mathbb{G}^{(\boldsymbol{\tau},\boldsymbol{X})}$ semimartingale. Since $\mathbb{N}^n \subset \mathbb{G}^{(\boldsymbol{\tau},\boldsymbol{X})}$ and M is adapted to $\mathbb{N}^n$, it follows from Theorem 5 that M is a $\mathbb{N}^n$ semimartingale. ■

1.3.3 Connection to filtration shrinkage

It has been observed that the optional projection of a local martingale M onto a filtration to which it is not adapted may lose the local martingale property, see Föllmer and Protter [50]. In that paper, the following general condition was given that guarantees that the local martingale property is preserved (see [50], Theorem 3.7):

Lemma 29 *Consider two filtrations $\mathbb{F} \subset \mathbb{G}$ and a $\mathbb{G}$ local martingale N. If there exists a sequence of $\mathbb{F}$ stopping times that reduce N, then its optional projection onto $\mathbb{F}$ is again a local martingale.*

Using this result we establish that certain local martingales that arise in the context of filtration expansion in fact retain their local martingale property when projected to various subfiltrations. In particular, if M is an $\mathbb{F}$ local martingale and a $\mathbb{G}$-semimartingale, then ${}^oM^{\mathbb{G}}$, the optional projection onto $\mathbb{F}$ of the local martingale part in the $\mathbb{G}$ semimartingale decomposition of M, always remains a local martingale.

Theorem 21 *Consider three filtrations $\mathbb{E} \subset \mathbb{F} \subset \mathbb{G}$, and let M be an $\mathbb{E}$ local martingale. Suppose M is also a $\mathbb{G}$ semimartingale with canonical decomposition $M = N + A$. Then the optional projection of N onto $\mathbb{F}$, when it exists, is again a local martingale.*

Remark. Note that M remains a special semimartingale in $\mathbb{G}$, given that it remains a semimartingale.

Proof. By Lemma 29 it suffices to show that N can be reduced using $\mathbb{E}$, and hence $\mathbb{F}$, stopping times. Now, since M is an $\mathbb{E}$ local martingale, it is locally in $\mathcal{H}^1_{loc}(\mathbb{E})$. Therefore it can be reduced with a sequene $(T_n)_{n\geq 1}$ of $\mathbb{E}$ stopping times such that for each n, $M^{T_n} \in \mathcal{H}^1(\mathbb{E})$. Also, M^{T_n} is a $\mathbb{G}$ semimartingale with canonical decomposition

$$M^{T_n} = N^{T_n} + A^{T_n}$$

A result by Yor ([128], théorème 5) then yields

$$\|N^{T_n}\|_{\mathcal{H}^1(\mathbb{G})} \leq c\|M^{T_n}\|_{\mathcal{H}^1(\mathbb{E})}$$

for some absolute constant c. Since M^{T_n} is in $\mathcal{H}^1(\mathbb{E})$, the right side is finite and it readily follows that N^{T_n} is a uniformly integrable $\mathbb{G}$ martingale for each n, which was what we set out to prove. ■

At this point, it is worth mentioning that the result above helps illustrate the role of our assumptions and shows how they fail to hold for the explicit counterexample given in [50]. There the authors consider a standard three-dimensional Brownian motion $(B_t)_{t\geq 0} = (B^1_t, B^2_t, B^3_t)_{t\geq 0}$ starting at $x_0 = (1, 0, 0)$. Let $\mathbb{H}$ be its natural filtration. It is well known that $(\|B_t\|)_{t\geq 0}$ is a Bessel(3) process with initial value 1 and that its reciprocal $N_t = \|B_t\|^{-1}$ is an $\mathbb{H}$ strict local martingale. Let $\mathbb{F}$ be the natural filtration of B^1. It is proved in [50] that the optional projection of N onto $\mathbb{F}$ is not a local martingale. Theorem 21 therefore implies that N can never appear as the local martingale part in the $\mathbb{H}$ decomposition $M = N + A$ of any $\mathbb{F}$ local martingale M that is also an $\mathbb{H}$ semimartingale.

As an application of Theorem 21 we obtain the following. The filtrations $\mathbb{F}$, $\mathbb{G}$ and $\mathbb{H}$ are now as in Theorem 16.

Corollary 9 *Let M be an $\mathbb{F}$ local martingale and let $\mathbb{H}$ be a filtration that coincides with $\mathbb{G}$ after τ. Suppose M is an $\mathbb{H}$ semimartingale with decomposition*

$$M_t = M_t^{\mathbb{H}} + A_t^{\mathbb{H}},$$

and suppose $A^{\mathbb{H}}$ has a $\mathbb{G}$ optional projection ${}^oA^{\mathbb{H}}$ which is locally integrable in $\mathbb{G}$. [1]

Then M is a $\mathbb{G}$ semimartingale with decomposition $M = M^{\mathbb{G}} + A^{\mathbb{G}}$, and $A^{\mathbb{G}} = (A^{\mathbb{H}})^p$. In particular, it follows that

$$(A^{\mathbb{H}})_t^p = \int_0^{t\wedge\tau} \frac{d\langle M, \mu\rangle_s + dJ_s}{Z_{s-}} + \int_{t\wedge\tau}^t dA_s^{\mathbb{H}}.$$

Here, the quantities Z, μ and J are defined as in Theorem 7.

Proof. M is an $\mathbb{H}$ semimartingale, hence a $\mathbb{G}$ semimartingale by Stricker's theorem [123]. Since ${}^oA^{\mathbb{H}}$ exists and M is $\mathbb{G}$ adapted, ${}^oM^{\mathbb{H}}$ exists. Taking optional projections of the relation $M = M^{\mathbb{H}} + A^{\mathbb{H}}$ and adding and subtracting $(A^{\mathbb{H}})^p$ yields

$$M = \left[{}^oM^{\mathbb{H}} + {}^oA^{\mathbb{H}} - (A^{\mathbb{H}})^p\right] + (A^{\mathbb{H}})^p.$$

By Theorem 21, ${}^oM^{\mathbb{H}}$ is a $\mathbb{G}$ local martingale, and therefore the quantity in brackets is a $\mathbb{G}$ local martingale. We deduce that $A^{\mathbb{G}} = (A^{\mathbb{H}})^p$. The expression for $(A^{\mathbb{H}})^p$ now follows from Theorem 16. ■

As a particular case, we recover Proposition 3.3 in [17]. There $\mathbb{G}$ and $\mathbb{H}$ are the progressive, respectively initial, expansions of $\mathbb{F}$ with a random time τ that avoids all $\mathbb{F}$ stopping times, and whose $\mathbb{F}$ conditional probabilities are equivalent to some deterministic measure. Under these assumptions, the authors compute explicitly the $\mathbb{G}$ dual predictable projection of $A_t^{\mathbb{H}} = \int_0^t k_s(\tau)ds$, where k is as in Theorem 6, and prove that

$$(A^{\mathbb{H}})_t^p = \int_0^{t\wedge\tau} \frac{d\langle M, Z\rangle_s}{Z_{s-}} + \int_{t\wedge\tau}^t k_s(\tau)ds.$$

This follows from Corollary 9 since $J = 0$ when τ avoids all $\mathbb{F}$ stopping times.

1.4 Filtration expansion with processes

In this section we study progressive filtration expansions with càdlàg processes. Using results from the weak convergence of σ-fields theory, we first establish a semimartingale

[1] This assumption guarantees that the dual predictable projection $(A^{\mathbb{H}})^p$ exists, and that ${}^oA^{\mathbb{H}} - (A^{\mathbb{H}})^p$ is a $\mathbb{G}$ local martingale, see Theorem VI.80 in [36].

convergence theorem, see Theorem 22 below. Then we apply it in a filtration expansion with a process setting and provide sufficient conditions for a semimartingale of the base filtration to remain a semimartingale in the expanded filtration. These main results are Theorems 25 and 27. Finally, an application to the expansion of a Brownian filtration with a time reversed diffusion is given through a detailed study and Kohatsu-Higa's example is recovered and generalized.

1.4.1 A semimartingale convergence result and applications to filtration expansion

In preparation for treating the expansion of filtrations via processes, we need to establish a general result on the convergence of semimartingales, which is definitely of interest in its own right.

1.4.1.1 A semimartingale convergence theorem

The following theorem is a generalization of the main result in [11].

Theorem 22 *Let $(\mathbb{G}^n)_{n\geq 1}$ be a sequence of right-continuous filtrations and let $\mathbb{G}$ be a filtration such that $\mathcal{G}^n_t \overset{w}{\to} \mathcal{G}_t$ for all t. Let $(X^n)_{n\geq 1}$ be a sequence of $\mathbb{G}^n$ semimartingales with canonical decomposition $X^n = X^n_0 + M^n + A^n$. Assume there exists $K > 0$ such that for all n,*

$$E(\int_0^T |dA^n_s|) \leq K \qquad \text{and} \qquad E(\sup_{0\leq s\leq T} |M^n_s|) \leq K$$

Then the following holds.

(i) Assume there exists a $\mathbb{G}$ adapted process X such that $E(\sup_{0\leq s\leq T} |X^n_s - X_s|) \to 0$. Then X is a $\mathbb{G}$ special semimartingale.

(ii) Moreover, assume $\mathbb{G}$ is right-continuous and let $X = M + A$ be the canonical decomposition of X. Then M is a $\mathbb{G}$ martingale and $\int_0^T |dA_s|$ and $\sup_{0\leq s\leq T} |M_s|$ are integrable.

Proof. Part (i). The idea of the proof of Part (i) is similar to the one in [11]. First, X is càdlàg since it is the a.s. uniform limit of a subsequence of the càdlàg processes $(X^n)_{n\geq 1}$. Also, since $||X^n_0 - X_0||_1 \to 0$, we can take w.l.o.g $X^n_0 = X_0 = 0$, and we do so. The integrability assumptions guarantee that $E(\sup_s |X^n_s|) \leq 2K$ and up to replacing K by $2K$, we assume that $E(\sup_s |M^n_s|) \leq K$, $E(\int_0^T |dA^n_s|) \leq K$ and $E(\sup_s |X^n_s|) \leq K$. Then $E(\sup_s |X_s|) \leq E(\sup_s |X_s - X^n_s|) + K$ and by taking limits $E(\sup_s |X_s|) \leq K$.

Let H be a $\mathbb{G}$ predictable elementary process of the form $H_t = \sum_{i=1}^k h_i 1_{]t_i,t_{i+1}]}(t)$, where h_i is a $\mathcal{G}_{t_i}$ measurable random variable such that $|h_i| \leq 1$ and $t_1 < \ldots < t_k < t_{k+1} = T$. Define now $H^n_t = \sum_{i=1}^k h^n_i 1_{]t_i,t_{i+1}]}(t)$, where $h^n_i = E(h_i \mid \mathcal{G}^n_{t_i})$. Then h^n_i is a $\mathcal{G}^n_{t_i}$ measurable random variable satisfying $|h_i^n| \leq 1$, hence H^n is a bounded $\mathbb{G}^n$ predictable elementary process. It follows that $H^n \cdot M^n$ is a $\mathbb{G}^n$ martingale and for each n,

$$|E((H^n \cdot X^n)_T)| \leq |E(\int_0^T H_s^n dA_s^n)| \leq E(\int_0^T |dA_s^n|) \leq K$$

Therefore, for each n,

$$|E((H \cdot X)_T)| \leq |E((H \cdot X)_T - (H^n \cdot X^n)_T)| + K \tag{1.5}$$

Since h_i is $\mathcal{G}_{t_i}$ measurable and $\mathcal{G}_t^n \xrightarrow{w} \mathcal{G}_t$ for all t, $h_i^n \xrightarrow{P} h_i$ for all $1 \leq i \leq k$. Since the set $\{1, \ldots, k\}$ is finite, successive extractions allow us to find a subsequence $\psi(n)$ (independent from i) such that for each $1 \leq i \leq k$, $h_i^{\psi(n)}$ converges a.s. to h_i. So up to working with the $\mathbb{G}^{\psi(n)}$ predictable elementary processes $H^{\psi(n)}$ and the stochastic integrals $H^{\psi(n)} \cdot X^{\psi(n)}$ in (1.5), we can assume that h_i^n converges a.s. to h_i, for each $1 \leq i \leq k$.

Now, $|E((H \cdot X)_T - (H^n \cdot X^n)_T)| \leq \sum_{i=1}^k E(|h_i Y_i - h_i^n Y_i^n|)$ where $Y_i = X_{t_{i+1}} - X_{t_i}$ and $Y_i^n = X_{t_{i+1}}^n - X_{t_i}^n$. Each term in the sum can be bounded as follows.

$$\begin{aligned}
E(|h_i Y_i - h_i^n Y_i^n|) &\leq E(|Y_i^n (h_i^n - h_i)|) + E(|h_i (Y_i^n - Y_i)|) \\
&\leq 2E(\sup_s |X_s^n||h_i^n - h_i|) + E(|Y_i^n - Y_i|) \\
&\leq 2E(\sup_s |X_s^n - X_s||h_i^n - h_i|) + 2E(\sup_s |X_s||h_i^n - h_i|) + 2E(\sup_s |X_s^n - X_s|) \\
&\leq 6E(\sup_s |X_s^n - X_s|) + 2E(\sup_s |X_s||h_i^n - h_i|)
\end{aligned}$$

Since $\sup_s |X_s||h_i^n - h_i|$ converges a.s to zero and that for all n, $|h_i^n| \leq 1$, hence $\sup_s |X_s||h_i^n - h_i| \leq 2 \sup_s |X_s|$ and $\sup_s |X_s| \in L^1$, the Dominated Convergence Theorem implies that $E(\sup_s |X_s||h_i^n - h_i|) \to 0$. Since by assumption $E(\sup_s |X_s^n - X_s|) \to 0$, it follows that $|E((H \cdot X)_T - (H^n \cdot X^n)_T)|$ converges to 0. Letting n tend to infinity in (1.5) gives $|E((H \cdot X)_T)| \leq K$. So X is a $\mathbb{G}$ quasimartingale, hence a $\mathbb{G}$ special semimartingale. Therefore X has a $\mathbb{G}$ canonical decomposition $X = M + A$ where M is a $\mathbb{G}$ local martingale and A is a $\mathbb{G}$ predictable finite variation process.

Part(ii). Let $(\tau_m)_{m \geq 1}$ be a sequence of bounded $\mathbb{G}$ stopping times that reduces M. Since for all t, $\mathcal{G}_t^m \xrightarrow{w} \mathcal{G}_t$, it follows from Lemma 20 that for each m there exist a function ϕ_m strictly increasing and a sequence $(\tau_m^n)_{n \geq 1}$ such that $(\tau_m^{\phi_m(n)})_{n \geq 1}$ converges in probability to τ_m and $\tau_m^{\phi_m(n)}$ are bounded $\mathbb{G}^{\phi_m(n)}$ stopping times. We can extract a subsequence $(\tau_m^{\phi_m(\psi_m(n))})_{n \geq 1}$ converging a.s. to τ_m. In order to simplify the notation, fix $m \geq 1$ and up to working with $\tilde{\mathbb{G}}^n = \mathbb{G}^{\phi_m(\psi_m(n))}$ instead of $\mathbb{G}^n$ (which satisfies the same assumptions), take $\Phi_m := \phi_m \circ \psi_m$ to be the identity. Let H be a $\mathbb{G}$ elementary predictable process as defined in Part (i). Since τ_m reduces M, $E((H \cdot A)_{\tau_m}) = E((H \cdot X)_{\tau_m})$. We can write

$$E((H \cdot A)_{\tau_m}) = E\big((H \cdot X)_{\tau_m} - (H^n \cdot X^n)_{\tau_m}\big) + E\big((H^n \cdot X^n)_{\tau_m} - (H^n \cdot X^n)_{\tau_m^n}\big) + E\big((H^n \cdot X^n)_{\tau_m^n}\big)$$

We start with the third term. Since $H^n \cdot M^n$ is a $\mathbb{G}^n$ martingale and τ_m^n is a bounded $\mathbb{G}^n$ stopping time, it follows from Doob's optional sampling theorem that $E((H^n \cdot X^n)_{\tau_m^n}) = E((H^n \cdot A^n)_{\tau_m^n})$, hence $|E((H^n \cdot X^n)_{\tau_m^n})| \leq E(\int_0^{\tau_m} |dA_s^n|) \leq K$.

We focus now on the first term. Let $Y^i_s = X_s - X_{t_i}$ and $Y^{i,n}_s = X^n_s - X^n_{t_i}$.

$$\begin{aligned}
E_1 &:= |E\big((H\cdot X)_{\tau_m} - (H^n \cdot X^n)_{\tau_m}\big)| \le E\Big(\sup_{0\le s\le T}\big|(H\cdot X)_s - (H^n\cdot X^n)_s\big|\Big)\\
&\le E\Big(\sum_{i=1}^k \sup_{t_i<s\le t_{i+1}}\big|h_i Y^i_s - h^n_i Y^{i,n}_s\big|\Big)\\
&\le \sum_{i=1}^k E\Big(\sup_{t_i<s\le t_{i+1}}|Y^{i,n}_s||h^n_i - h_i| + \sup_{t_i<s\le t_{i+1}}|h_i||Y^{i,n}_s - Y^i_s|\Big)
\end{aligned}$$

Since $|h^i|\le 1$, $|Y^{i,n}_s| \le 2\sup_u |X^n_u|$ and $|Y^{i,n}_s - Y^i_s| \le 2\sup_u |X^n_u - X_u|$, it follows that

$$\begin{aligned}
E_1 &\le 2\sum_{i=1}^k \big\{E\big(\sup_{0\le u\le T}|X^n_s||h^n_i - h_i|\big) + E\big(\sup_{0\le u\le T}|X^n_u - X_u|\big)\big\}\\
&\le 6kE(\sup_{0\le s\le T}|X^n_s - X_s|) + 2\sum_{i=1}^k\{E\big(\sup_{0\le s\le T}|X_s||h^n_i - h_i|\big)\}
\end{aligned}$$

We study now the second term $E_2 := E\big((H^n\cdot X^n)_{\tau_m} - (H^n\cdot X^n)_{\tau^n_m}\big)$. Let $0<\eta<\min_i|t_{i+1}-t_i|$, and define $Y^n = H^n\cdot X^n$. Write now

$$E\big(|Y^n_{\tau_m} - Y^n_{\tau^n_m}|\big) = E\big(|Y^n_{\tau_m} - Y^n_{\tau^n_m}|1_{\{|\tau_m-\tau^n_m|\le\eta\}}\big) + E\big(|Y^n_{\tau_m} - Y^n_{\tau^n_m}|1_{\{|\tau_m-\tau^n_m|>\eta\}}\big) =: e_1 + e_2$$

We study each of the two terms separately. We start with e_2.

$$\begin{aligned}
e_2 &\le E\big((|Y^n_{\tau_m}| + |Y^n_{\tau^n_m}|)1_{\{|\tau_m-\tau^n_m|>\eta\}}\big) \le 2E\Big(\sum_{i=1}^k \sup_{t_i<s\le t_{i+1}}|X^n_s - X^n_{t_i}|1_{\{|\tau_m-\tau^n_m|>\eta\}}\Big)\\
&\le 4kE\Big(\sup_{0\le u\le T}|X^n_u|1_{\{|\tau_m-\tau^n_m|>\eta\}}\Big)\\
&\le 4kE\Big(\sup_{0\le u\le T}|X^n_u - X_u|\Big) + 4kE\Big(\sup_{0\le u\le T}|X_u|1_{\{|\tau_m-\tau^n_m|>\eta\}}\Big)
\end{aligned}$$

We study now e_1. On $\{|\tau_m - \tau^n_m|\le\eta\}$ and since $\eta<\min_i|t_{i+1}-t_i|$, we have $|Y^n_{\tau_m} - Y^n_{\tau^n_m}| \le 2\sup_{s\le t\le s+\eta}|X^n_t - X^n_s|$. In fact, one of the two following cases is possible for τ_m and τ^n_m. Either they are both in the same interval $(t_i, t_{i+1}]$, in which case, $|Y^n_{\tau_m} - Y^n_{\tau^n_m}| = |h^n_i(X^n_{\tau_m} - X^n_{\tau^n_m})| \le \sup_{s\le t\le s+\eta}|X^n_t - X^n_s|$, or they are in two consecutive intervals. For the second case, take for example $t_{i-1}<\tau_m\le t_i<\tau^n_m\le t_{i+1}$, then

$$\begin{aligned}
|Y^n_{\tau_m} - Y^n_{\tau^n_m}| &= |h^n_{i-1}(X^n_{\tau_m} - X^n_{t_{i-1}}) - h^n_{i-1}(X^n_{t_i} - X^n_{t_{i-1}}) - h^n_i(X^n_{\tau^n_m} - X^n_{t_i})|\\
&= |h^n_{i-1}(X^n_{\tau_m} - X^n_{t_i}) - h^n_i(X^n_{\tau^n_m} - X^n_{t_i})| \le |X^n_{\tau_m} - X^n_{t_i}| + |X^n_{\tau^n_m} - X^n_{t_i}|\\
&\le 2\sup_{s\le t\le s+\eta}|X^n_t - X^n_s|
\end{aligned}$$

The case $t_{i-1}<\tau^n_m\le t_i<\tau_m\le t_{i+1}$ is similar. Hence

$$e_1 \le 2E(\sup_{s\le t\le s+\eta}|X^n_t - X^n_s|1_{\{|\tau_m-\tau^n_m|\le\eta\}}) \le 2E(\sup_{s\le t\le s+\eta}|X^n_t - X^n_s|)$$

Putting all this together yields for each $0 < \eta < \min_i |t_{i+1} - t_i|$ and each n

$$|E(H \cdot A)_{\tau_m}| \leq K + 2E(\sup_{s \leq t \leq s+\eta} |X_t^n - X_s^n|) + 2\sum_{i=1}^{k} E(\sup_{0 \leq s \leq T} |X_s||h_i^n - h_i|) \\ + 4kE(\sup_{0 \leq s \leq T} |X_s| 1_{\{|\tau_m - \tau_m^n| > \eta\}}) + 10kE\big(\sup_{0 \leq u \leq T} |X_u^n - X_u|\big)$$

Getting back to the general case, we obtain for each $m \geq 1$ and each $n \geq 1$,

$$|E(H \cdot A)_{\tau_m}| \leq K + 2E(\sup_{s \leq t \leq s+\eta} |X_t^{\Phi_m(n)} - X_s^{\Phi_m(n)}|) + 2\sum_{i=1}^{k} E(\sup_{0 \leq s \leq T} |X_s||h_i^{\Phi_m(n)} - h_i|) \\ + 4kE(\sup_{0 \leq s \leq T} |X_s| 1_{\{|\tau_m - \tau_m^{\Phi_m(n)}| > \eta\}}) + 10kE\big(\sup_{0 \leq u \leq T} |X_u^{\Phi_m(n)} - X_u|\big)$$

As in the proof of Part (i), successive extractions allow us to find $\lambda_m(n)$ (independent from i) such that for all $1 \leq i \leq k$, $h_i^{\lambda_m(n)}$ converges a.s. to h_i. Letting n go to infinity in

$$|E(H \cdot A)_{\tau_m}| \leq K + 2E(\sup_{s \leq t \leq s+\eta} |X_t^{\lambda_m(n)} - X_s^{\lambda_m(n)}|) + 2\sum_{i=1}^{k} E(\sup_{0 \leq s \leq T} |X_s||h_i^{\lambda_m(n)} - h_i|) \\ + 4kE(\sup_{0 \leq s \leq T} |X_s| 1_{\{|\tau_m - \tau_m^{\lambda_m(n)}| > \eta\}}) + 10kE\big(\sup_{0 \leq u \leq T} |X_u^{\lambda_m(n)} - X_u|\big)$$

gives the estimate

$$|E(H \cdot A)_{\tau_m}| \leq K + 2\limsup_{n \to \infty} E(\sup_{s \leq t \leq s+\eta} |X_t^n - X_s^n|)$$

Let $\lim_{\eta \to 0} \limsup_{n \to \infty} E(\sup_{s \leq t \leq s+\eta} |X_s^n - X_t^n|) = C$. Since $E(\sup_{s \leq t \leq s+\eta} |X_t^n - X_s^n|) \leq 2E(\sup_u |X_u^n|) \leq 2(E(\sup_u |M_u^n| + \int_0^T |dA_s^n|)) \leq 4K$, $C < \infty$. Now letting η go to zero yields finally $|E(H \cdot A)_{\tau_m}| \leq K + 2C$, for each m. Thus $E(\int_0^{\tau_m} |dA_s|) \leq K + 2C$, for each m and hence $E(\int_0^T |dA_s|) \leq K + 2C$.

Now, $M = X - A = (X - X^n) + M^n + A^n - A$, and so

$$\sup_{0 \leq s \leq T} |M_s| \leq \sup_{0 \leq s \leq T} |X_s - X_s^n| + \sup_{0 \leq s \leq T} |M^n| + \int_0^T |dA_s^n| + \int_0^T |dA_s|$$

Thus $E(\sup_{0 \leq s \leq T} |M_s|) \leq 3K + 2C$ and M is a $\mathbb{G}$ martingale. ■

Once one obtains that X is a $\mathbb{G}$ special semimartingale, one can be interested in characterizing the martingale M and the finite variation predictable process A in terms of the processes M^n and A^n. Mémin ([111, Theorem 11]) achieved this under *extended convergence*. Recall the following definition.

Definition 8 *We say that $(X^n, \mathbb{G}^n)$ converges to $(X, \mathbb{G})$ in the extended sense if for every $G \in \mathbb{G}_T$, the sequence of càdlàg processes $(X_t^n, E(1_G \mid \mathcal{G}_t^n))_{0 \leq t \leq T}$ converges in probability under the Skorohod J_1 topology to $(X_t, E(1_G \mid \mathcal{G}_t))_{0 \leq t \leq T}$.*

The author proves the following theorem in [111] to which we refer for a proof. In the theorem below, $\mathbb{G}^n$ and $\mathbb{G}$ are right-continuous filtrations.

Theorem 23 *Let $(X^n)_{n\geq 1}$ be a sequence of $\mathbb{G}^n$ special semimartingales with canonical decompositions $X^n = M^n + A^n$ where M^n is a $\mathbb{G}^n$ martingale and A^n is a $\mathbb{G}^n$ predictable finite variation process. We suppose that the sequence $([X^n, X^n]_T^{\frac{1}{2}})_{n\geq 1}$ is uniformly integrable and that the sequence $(V(A^n)_T)_{n\geq 1}$ (where V denotes the variation process) of real random variables is tight in $\mathbb{R}$. Let X be a $\mathbb{G}$ quasi-left continuous special semimartingale with a canonical decomposition $X = M + A$ such that $([X, X]_T^{\frac{1}{2}}) < \infty$.*

If the extended convergence $(X^n, \mathbb{G}^n) \to (X, \mathbb{G})$ holds, then (X^n, M^n, A^n) converges in probability under the Skorohod J_1 topology to (X, M, A).

In a filtration expansion setting, the sequence X^n is constant and equal to some semimartingale X of the base filtration. In this case the extended convergence assumption in Theorem 23 reduces to the weak convergence of the filtrations. We can deduce the following corollary from Theorems 22 and 23.

Corollary 10 *Let $(\mathbb{G}^n)_{n\geq 1}$ be a sequence of right-continuous filtrations and let $\mathbb{G}$ be a filtration such that $\mathcal{G}_t^n \xrightarrow{w} \mathcal{G}_t$ for all t. Let X be a stochastic process such that for each n, X is a $\mathbb{G}^n$ semimartingale with canonical decomposition $X = M^n + A^n$ such that there exists $K > 0$, $E(\int_0^T |dA_s^n|) \leq K$ and $E(\sup_{0\leq s\leq T} |M_s^n|) \leq K$ for all n. Then*

(i) If X is $\mathbb{G}$ adapted, then X is a $\mathbb{G}$ special semimartingale.

(ii) Assume moreover that $\mathbb{G}$ is right-continuous and let $X = M + A$ be the canonical decomposition of X. Then M is a $\mathbb{G}$ martingale and $\sup_{0\leq s\leq T} |M_s|$ and $\int_0^T |dA_s|$ are integrable.

(iii) Furthermore, assume that X is $\mathbb{G}$ quasi-left continuous and $\mathbb{G}^n \xrightarrow{w} \mathbb{G}$. Then (M^n, A^n) converges in probability under the Skorohod J_1 topology to (M, A).

Proof. The sequence $X^n = X$ clearly satisfies the assumptions of Theorem 22, and the two first claims follow. For the last claim, notice that $[X, X]_T \in L^1$, so $\sqrt{[X, X]_T} \in L^1$ and hence $(\sqrt{[X, X]_t})_{0\leq t\leq T}$ is a uniformly integrable family of random variables. The tightness of the sequence of random variables $(V(A^n)_T)_{n\geq 1}$ follows from $E(\int_0^T |dA_s^n|) \leq K$ for any n and some K independent from n. ■

We provide in the next subsection a first application to the initial and progressive filtration expansions with a random variable and a general theorem on the progressive expansion with a process. We assume in the sequel that a right-continuous filtration $\mathbb{F}$ is given.

1.4.1.2 Application to initial and progressive filtration expansions with a random variable

Assume that $\mathbb{F}$ is the natural filtration of some càdlàg process. Let τ be a random variable and $\mathbb{H}$ and $\mathbb{G}$ the initial and progressive expansions of $\mathbb{F}$ with τ. In this subsection, the filtration $\mathbb{G}$ is considered only when τ is non negative. It is proved in section 1.3 that if τ satisfies Jacod's criterion i.e. if there exists a σ-finite measure η on $\mathcal{B}(\mathbb{R})$ such that

$$P(\tau \in \cdot \mid \mathcal{F}_t)(\omega) \ll \eta(.) \quad \text{a.s.}$$

then every $\mathbb{F}$ semimartingale remains an $\mathbb{H}$ and $\mathbb{G}$ semimartingale. That it is an $\mathbb{H}$ semimartingale is due to Jacod's theorem. That it is also a $\mathbb{G}$ semimartingale follows from Theorem 5. Its $\mathbb{G}$ decomposition is obtained in the previous section and this relies on the fact that these two filtrations *coincide* after τ. We provide now a similar but partial result for a random variable τ which may not satisfy Jacod's criterion. Assume there exists a sequence of random times $(\tau_n)_{n\geq 0}$ converging in probability to τ and let $\mathbb{H}^n$ be the initial expansion of $\mathbb{F}$ with τ_n. The following holds.

Theorem 24 *Let M be an $\mathbb{F}$ martingale such that $\sup_{0\leq t\leq T}|M_t|$ is integrable. Assume there exists an $\mathbb{H}^n$ predictable finite variation process A^n such that $M - A^n$ is an $\mathbb{H}^n$ martingale. If there exists K such that $E(\int_0^T |dA^n_s|) \leq K$ for all n, then M is an $\mathbb{H}$ and $\mathbb{G}$ semimartingale.*

Proof. Since τ_n converges in probability to τ and $\mathbb{F}$ is the natural filtration of some càdlàg process, we can prove that $\mathcal{H}^n_t \overset{w}{\to} \mathcal{H}_t$ for each $t \in [0,T]$, using the same techniques as in Lemmas 15 and 16. Up to replacing K by $K + E(\sup_{0\leq t\leq T}|M_t|)$, $M^n = M - A^n$ and A^n satisfy the assumptions of Corollary 10. Therefore M is an $\mathbb{H}$ semimartingale, and a $\mathbb{G}$ semimartingale by Stricker's theorem. ■

One case where the first assumption of Theorem 24 is satisfied is when τ_n satisfies Jacod's criterion, for each $n \geq 0$. In this case, and if $\mathbb{F}$ is the natural filtration of a Brownian motion W, the result above can be made more explicit. Assume for simplicity that the conditional distributions of τ_n are absolutely continuous w.r.t Lebesgue measure,

$$P(\tau_n \in du \mid \mathcal{F}_t)(\omega) = p^n_t(u,\omega)du$$

where the conditional densities are chosen so that $(u,\omega,t) \to p^n_t(u,\omega)$ is càdlàg in t and measurable for the optional σ-field associated with the filtration $\hat{\mathbb{F}}$ given by $\hat{\mathcal{F}}_t = \cap_{u>t}\mathcal{B}(\mathbb{R}) \otimes \mathcal{F}_u$. From the martingale representation theorem in a Brownian filtration, there exists for each n a family $\{q^n(u), u > 0\}$ of $\mathbb{F}$ predictable processes $(q^n_t(u))_{0\leq t\leq T}$ such that

$$p^n_t(u) = p^n_0(u) + \int_0^t q^n_s(u)dW_s \tag{1.6}$$

Corollary 11 *Assume there exists K such that $E\Big(\int_0^T \Big|\frac{q^n_s(\tau_n)}{p^n_s(\tau_n)}\Big|ds\Big) \leq K$ for all n, then W is a special semimartingale in both $\mathbb{H}$ and $\mathbb{G}$.*

Proof. Since τ_n satisfy Jacod's criterion, it follows from Theorem 2.1 in [72] that $W_t - A^n_t$ is an $\mathbb{H}$ local martingale, where $A^n_t = \int_0^t \frac{d\langle p^n(u), W\rangle_s}{p^n_s(u)}\Big|_{u=\tau_n}$. Now, it follows from (1.6) that $A^n_t = \int_0^t \frac{q^n_s(\tau_n)}{p^n_s(\tau_n)} ds$ and Theorem 24 allows us to conclude. ∎

Assume the assumptions of Corollary 11 are satisfied and let $W = M + A$ be the $\mathbb{H}$ canonical decomposition of W. Let m be an $\mathbb{F}$ predictable process such that $\int_0^t m_s^2 ds$ is locally integrable and let M be the $\mathbb{F}$ local martingale $M_t = \int_0^t m_s dW_s$. Theorem VI.5 in [117] then guarantees that M is an $\mathbb{H}$ semimartingale as soon as the process $(\int_0^t m_s dA_s)_{t\geq 0}$ exists as a path-by-path Lebesgue-Stieltjes integral a.s.

Example 4 *In order to emphasize that some assumptions as in Theorem 24 are needed, we provide now an example where the conclusion of Theorem 24 fails to hold. Let $\mathbb{F}$ be the natural filtration of some Brownian motion B and choose τ to be some functional of the Brownian path i.e. $\tau = f((B_s, 0 \leq s \leq 1))$, such that $\sigma(\tau) = \mathcal{F}_1$. Then B is not a semimartingale in $\mathbb{H} = (\mathcal{F}_t \vee \sigma(\tau))_{0\leq t\leq 1}$. Now, define $\tau_n = \tau + \frac{1}{\sqrt{n}} N$, where N is a standard normal random variable independent from $\mathbb{F}$. Then τ_n converge a.s. to τ and $P(\tau_n \leq u \mid \mathcal{F}_t) = \int_{-\infty}^u E(g_n(v-\tau) \mid \mathcal{F}_t) dv$ where g_n is the probability density function of $\frac{1}{\sqrt{n}} N$, hence $P(\tau_n \in du \mid \mathcal{F}_t)(\omega) = E(g_n(u-\tau) \mid \mathcal{F}_t)(\omega)(du)$. Therefore, τ_n satisfies Jacod's criterion, for each n and $p^n_t(u,\omega) = E(g_n(u-\tau) \mid \mathcal{F}_t)(\omega)$. Thus, B is a semimartingale in $\mathbb{H}^n = (\mathcal{F}_t \vee \sigma(\tau_n)_{0\leq t\leq 1})$ and $\mathcal{H}^n_t \xrightarrow{w} \mathcal{H}_t$ for each $0 \leq t \leq 1$.*

1.4.1.3 Application to filtration expansion with a process

Let $(N^n)_{n\geq 1}$ be a sequence of càdlàg processes converging in probability under the Skorohod J_1-topology to a càdlàg process N and let $\mathbb{N}^n$ and $\mathbb{N}$ be their natural filtrations. Define the filtrations $\mathbb{G}^{0,n} = \mathbb{F} \vee \mathbb{N}^n$ and $\mathbb{G}^n$ by $\mathcal{G}^n_t = \bigcap_{u>t} \mathcal{G}^{0,n}_u$. Let also $\mathbb{G}^0$ (resp. $\mathbb{G}$) be the smallest (resp. the smallest right-continuous) filtration containing $\mathbb{F}$ and to which N is adapted. The result below is the main theorem of this section.

Theorem 25 *Let X be an $\mathbb{F}$ semimartingale such that for each n, X is a $\mathbb{G}^n$ semimartingale with canonical decomposition $X = M^n + A^n$. Assume $E(\int_0^T |dA^n_s|) \leq K$ and $E(\sup_{0\leq s\leq T} |M^n_s|) \leq K$ for some K and all n. Finally, assume one of the following holds.*

- *N has no fixed times of discontinuity,*
- *N^n is a discretization of N along some refining subdivision $(\pi_n)_{n\geq 1}$ such that each fixed time of discontinuity of N belongs to $\cap_n \pi_n$.*

Then

(i) X is a $\mathbb{G}^0$ special semimartingale.

(ii) Moreover, if $\mathbb{F}$ is the natural filtration of some càdlàg process then X is a $\mathbb{G}$ special semimartingale with canonical decomposition $X = M + A$ such that M is a $\mathbb{G}$ martingale and $\sup_{0\leq s\leq T} |M_s|$ and $\int_0^T |dA_s|$ are integrable.

(iii) *Furthermore, assume that X is $\mathbb{G}$ quasi-left continuous and $\mathbb{G}^n \xrightarrow{w} \mathbb{G}$. Then (M^n, A^n) converges in probability under the Skorohod J_1 topology to (M, A).*

Proof. Under assumption (i), since $N^n \xrightarrow{P} N$ and $P(\Delta N_t \neq 0) = 0$ for all t, it follows from Lemma 16 that $\mathcal{N}_t^n \xrightarrow{P} \mathcal{N}_t$ for all t. The same holds under assumption (ii) using Lemma 17. Lemma 18 then ensures that $\mathcal{G}_t^{0,n} \xrightarrow{w} \mathcal{G}_t^0$ for all t. Since $\mathcal{G}_t^{0,n} \subset \bigcap_{u>t} \mathcal{G}_u^{0,n} = \mathcal{G}_t^n$, it follows from Lemma 19 that $\mathcal{G}_t^n \xrightarrow{w} \mathcal{G}_t^0$ for all t. Being an $\mathbb{F}$ semimartingale, X is clearly $\mathbb{G}^0$ adapted. An application of Corollary 10 ends the proof of the first claim. When $\mathbb{F}$ is the natural filtration of some càdlàg process, the same proofs as of Lemmas 16 and 17 guarantee that $\mathcal{G}_t^n \xrightarrow{w} \mathcal{G}_t$ for all t. Since $\mathbb{G}$ is right-continuous, the second and third claims follow from Corollary 10. ∎

We apply this result to expand the filtration $\mathbb{F}$ progressively with a point process. Let $(\tau_i)_{i\geq 1}$ and $(X_i)_{i\geq 1}$ be two sequences of random variables such that for each n, the random vector $(\tau_1, X_1, \dots, \tau_n, X_n)$ satisfies Jacod's criterion w.r.t the filtration $\mathbb{F}$. Assume that for all t and i, $P(\tau_i = t) = 0$ and that one of the following holds:

(i) For all i, X_i and τ_i are independent, $E|X_i| = \mu$ for some μ and $\sum_{i=1}^{\infty} P(\tau_i \leq T) < \infty$

(ii) $E(|X_i^2|) = c$ and $\sum_{i=1}^{\infty} \sqrt{P(\tau_i \leq T)} < \infty$.

Let $N_t^n = \sum_{i=1}^{n} X_i 1_{\{\tau_i \leq t\}}$ and $N_t = \sum_{i=1}^{\infty} X_i 1_{\{\tau_i \leq t\}}$. The assumptions on N^n and N as of Theorem 25 are satisfied.

Lemma 30 *Under the assumptions above, $N_t \in L^1$ for each t, $N^n \xrightarrow{P} N$ and N has no fixed times of discontinuity.*

Proof. We prove the statement under assumption (i). For each t,

$$E(|N_t|) \leq \sum_{i=1}^{\infty} E(|X_i| 1_{\{\tau_i \leq t\}}) \leq \mu \sum_{i=1}^{\infty} P(\tau_i \leq t) < \infty$$

Therefore, $N_t \in L^1$. For $\eta > 0$ and n integer, we obtain the following estimate.

$$\begin{aligned} P(\sup_{0\leq t\leq T} |N_t - N_t^n| \geq \eta) &= P(\sup_{0\leq t\leq T} \Big| \sum_{i=n+1}^{\infty} X_i 1_{\{\tau_i \leq t\}} \Big| \geq \eta) \\ &\leq P(\sup_{0\leq t\leq T} \sum_{i=n+1}^{\infty} |X_i| 1_{\{\tau_i \leq t\}} \geq \eta) = P(\sum_{i=n+1}^{\infty} |X_i| 1_{\{\tau_i \leq T\}} \geq \eta) \\ &\leq \frac{1}{\eta} E(\sum_{i=n+1}^{\infty} |X_i| 1_{\{\tau_i \leq T\}}) = \frac{\mu}{\eta} \sum_{i=n+1}^{\infty} P(\tau_i \leq T) \to 0 \end{aligned}$$

This implies $N^n \xrightarrow{P} N$. Under assumption (ii), the proof is also straightforward and based on Cauchy Schwarz inequalities. Finally, since

$$P(|\Delta N_t| \neq 0) \leq P(\exists i \mid \tau_i = t) \leq \sum_{i=1}^{\infty} P(\tau_i = t) = 0$$

N has no fixed times of discontinuity ■

Since the random vector $(\tau_1, X_1, \ldots, \tau_n, X_n)$ is assumed to satisfy Jacod's criterion, it follows from Corollary 8 that $\mathbb{F}$ semimartingales remain $\mathbb{G}^n$ semimartingales, for each n. Therefore, this property also holds between $\mathbb{F}$ and $\mathbb{G}$ for $\mathbb{F}$ semimartingales whose $\mathbb{G}^n$ canonical decompositions satisfy the regularity assumptions of Theorem 25. Here $\mathbb{G}$ is the smallest filtration containing $\mathbb{F}$ and to which N is adapted.

We would like to take a step further and reverse the previous situation. That is instead of starting with a sequence of processes N^n converging to some process N, and putting assumptions on the semimartingale properties of $\mathbb{F}$ semimartingales w.r.t the intermediate filtrations $\mathbb{G}^n$ and their decompositions therein, we would like to expand the filtration $\mathbb{F}$ with a given process X and express all the assumptions in terms of X and the $\mathbb{F}$ semimartingales considered. We are able to do this for càdlàg processes which satisfy a criterion that can loosely be seen as a localized extension of Jacod's criterion to processes. The integrability assumptions of Theorem 25 are expressed in terms of $\mathcal{F}_t$-conditional densities. Before doing this, we conclude this subsection by studying the stability of hypothesis (H) with respect to the weak convergence of the σ-fields in a filtration expansion setting.

1.4.1.4 The case of hypothesis (H)

Recall that given two nested filtrations $\mathbb{F} \subset \mathbb{G}$, we say that hypothesis (H) holds between $\mathbb{F}$ and $\mathbb{G}$ if any square integrable $\mathbb{F}$ martingale remains a $\mathbb{G}$ martingale. This hypothesis was studied by Dellacherie and Meyer in [35] and Brémaud and Yor, see [14] where a proof of the next Lemma 31 is available.

Lemma 31 *Let $\mathbb{F} \subset \mathbb{G}$ two nested filtrations. The following assertions are equivalent.*

(i) Hypothesis (H) holds between $\mathbb{F}$ and $\mathbb{G}$.

(ii) For each $0 \leq t \leq T$, $\mathcal{F}_T$ and $\mathcal{G}_t$ are conditionally independent given $\mathcal{F}_t$.

(iii) For each $0 \leq t \leq T$, each $F \in L^2(\mathcal{F}_T)$ and each $G_t \in L^2(\mathcal{G}_t)$,

$$E(FG_t \mid \mathcal{F}_t) = E(F \mid \mathcal{F}_t)E(G_t \mid \mathcal{F}_t).$$

Let $\mathbb{F} \subset \mathbb{G}$ be two nested right-continuous filtrations and $\mathbb{G}^n$ be a sequence of right-continuous filtrations containing $\mathbb{F}$ and such that $\mathcal{G}_t^n$ converges weakly to $\mathcal{G}_t$ for each t. We mentioned that an $\mathbb{F}$ local martingale that remains a $\mathbb{G}^n$ semimartingale for each n might still lose its semimartingale property in $\mathbb{G}$ and we provided conditions that prevent this pathological behavior. In this subsection, we prove that this cannot happen in case hypothesis (H) holds between $\mathbb{F}$ and each $\mathbb{G}^n$. One obtains even that hypothesis (H) holds between $\mathbb{F}$ and $\mathbb{G}$.

Theorem 26 *Let $\mathbb{F}$, $\mathbb{G}$ and $(\mathbb{G}^n)_{n\geq 1}$ right-continuous filtrations such that $\mathbb{F} \subset \mathbb{G}$, $\mathbb{F} \subset \mathbb{G}^n$ for each n and $\mathcal{G}_t^n \xrightarrow{w} \mathcal{G}_t$ for each t. Assume that for each n, hypothesis (H) holds between $\mathbb{F}$ and $\mathbb{G}^n$. Then hypothesis (H) holds between $\mathbb{F}$ and $\mathbb{G}$.*

Proof. We use Lemma 31 and start with the bounded case. Let $0 \leq t \leq T$, $F \in L^2(\mathcal{F}_T)$ and $G_t \in L^\infty(\mathcal{G}_t)$. For each n, define $G_t^n = E(G_t \mid \mathcal{G}_t^n)$. Then $G_t^n \in L^\infty(\mathcal{G}_t^n)$. Since hypothesis (H) holds between $\mathbb{F}$ and $\mathbb{G}^n$, Lemma 31 guarantees that $E(FG_t^n \mid \mathcal{F}_t) = E(F \mid \mathcal{F}_t)E(G_t^n \mid \mathcal{F}_t)$. But $\mathcal{F}_t \subset \mathcal{G}_t^n$, hence $E(G_t^n \mid \mathcal{F}_t) = E(E(G_t \mid \mathcal{G}_t^n) \mid \mathcal{F}_t) = E(G_t \mid \mathcal{F}_t)$. Since $\mathcal{G}_t^n \xrightarrow{w} \mathcal{G}_t$, $FG_t^n \xrightarrow{P} FG_t$. Now FG_t^n is bounded by a square integrable process (by assumption) so the convergence holds in L^1 by the Dominated Convergence theorem so that $E(FG_t^n \mid \mathcal{F}_t) \xrightarrow{P} E(FG_t \mid \mathcal{F}_t)$. This proves that $E(FG_t \mid \mathcal{F}_t) = E(F \mid \mathcal{F}_t)E(G_t \mid \mathcal{F}_t)$. The general case where $G_t \in L^2(\mathcal{G}_t)$ follows by applying the bounded case result to the bounded random variables $G_t^{(m)} = G_t \wedge m$. Then for each m,

$$E(FG_t^{(m)} \mid \mathcal{F}_t) \xrightarrow{P} E(FG_t \mid \mathcal{F}_t)$$

and the Monotone Convergence theorem allows us to conclude. ■

A similar result is straightforward and holds if we assume that the sequence of filtrations $(\mathbb{G}^n)_{n\geq 1}$ converges weakly to the filtration $\mathbb{G}$ rather than assuming the convergence of the corresponding σ-fields.

Lemma 32 *Assume $\mathbb{G}^n \xrightarrow{w} \mathbb{G}$. Let M be a $\mathbb{G}$ adapted process such that M is a $\mathbb{G}^n$ martingale for each $n \geq 1$. Then M is also a $\mathbb{G}$ martingale.*

Proof. Since $\mathbb{G}^n \xrightarrow{w} \mathbb{G}$ and M_T is $\mathcal{G}_T$ measurable, it follows that $E(M_T \mid \mathcal{G}_.^n) \to E(M_T \mid \mathcal{G}_.)$. Now since M is a $\mathbb{G}^n$ martingale, $M = E(M_T \mid \mathcal{G}^n)$ is independent from n, therefore $M = E(M_T \mid \mathcal{G})$. ■

We can now provide the following result whose proof is immediate given the previous lemma.

Corollary 12 *Let $\mathbb{F}$, $\mathbb{G}$ and $(\mathbb{G}^n)_{n\geq 1}$ right-continuous filtrations such that $\mathbb{F} \subset \mathbb{G}$, $\mathbb{F} \subset \mathbb{G}^n$ for each n and $\mathbb{G}^n \xrightarrow{w} \mathbb{G}$. Assume that for each n, hypothesis (H) holds between $\mathbb{F}$ and $\mathbb{G}^n$. Then hypothesis (H) holds between $\mathbb{F}$ and $\mathbb{G}$.*

1.4.2 Results based on Jacod's type criterion for the increments of the process

In this section, we assume a càdlàg process X and a right-continuous filtration $\mathbb{F}$ are given. We study the case where the process X and the filtration $\mathbb{F}$ satisfy the following assumption.

Assumption 2 (Generalized Jacod's criterion) *There exists a sequence $(\pi_n)_{n\geq 1} = (\{t_i^n\})_{n\geq 1}$ of subdivisions of $[0,T]$ whose mesh tends to zero and such that for each n, $(X_{t_0^n}, X_{t_1^n} - X_{t_0^n}, \ldots, X_T - X_{t_n^n})$ satisfies Jacod's criterion, i.e. there exists a σ-finite measure η_n on $\mathcal{B}(\mathbb{R}^{n+2})$ such that $P\big((X_{t_0^n}, X_{t_1^n} - X_{t_0^n}, \ldots, X_T - X_{t_n^n}) \in \cdot \mid \mathcal{F}_t\big)(\omega) \ll \eta_n(\cdot)$ a.s.*

Under Assumption 2, the $\mathcal{F}_t$-conditional density

$$p_t^{(n)}(u_0,\dots,u_{n+1},\omega)=\frac{P\big((X_{t_0^n},X_{t_1^n}-X_{t_0^n},\dots,X_T-X_{t_n^n})\in(du_0,\dots,du_{n+1})\mid\mathcal{F}_t\big)(\omega)}{\eta_n(du_0,\dots,du_{n+1})}$$

exists for each n, and can be chosen so that $(u_0,\dots,u_{n+1},\omega,t)\to p_t^{(n)}(u_0,\dots,u_{n+1},\omega)$ is càdlàg in t and measurable for the optional σ-field associated with the filtration $\hat{\mathbb{F}}_t$ given by $\hat{\mathcal{F}}_t=\cap_{u>t}\mathcal{B}(\mathbb{R}^{n+2})\otimes\mathcal{F}_u$. For each $0\le i\le n$, define

$$p_t^{i,n}(u_0,\dots,u_i)=\int_{\mathbb{R}^{n+1-i}}p_t^{(n)}(u_0,\dots,u_{n+1})\eta_n(du_{i+1},\dots,du_{n+1})$$

Let M be a continuous $\mathbb{F}$ local martingale. Define

$$A_t^{i,n}=\int_0^t\frac{d\langle p^{i,n}(u_0,\dots,u_i),M\rangle_s}{p_{s-}^{i,n}(u_0,\dots,u_i)}\Big|_{\forall 0\le k\le i,u_k=X_{t_k^n}-X_{t_{k-1}^n}}\tag{1.7}$$

Finally define

$$A_t^{(n)}=\sum_{i=0}^n\int_{t\wedge t_i^n}^{t\wedge t_{i+1}^n}dA_s^{i,n}$$

i.e.

$$A_t^{(n)}=\sum_{i=0}^n 1_{\{t_i^n\le t<t_{i+1}^n\}}\Big(\sum_{k=0}^{i-1}\int_{t_k^n}^{t_{k+1}^n}dA_s^{k,n}+\int_{t_i^n}^t dA_s^{i,n}\Big)\tag{1.8}$$

Of course, on each time interval $\{t_i^n\le t<t_{i+1}^n\}$, only one term appears in the outer sum. Let $\mathbb{G}^0$ (resp. $\mathbb{G}$) be the smallest (resp. the smallest right-continuous) filtration containing $\mathbb{F}$ and relative to which X is adapted. The theorem below is the main result of this section.

Theorem 27 *Assume X and $\mathbb{F}$ satisfy Assumption 2 and that one of the following holds.*

- *X has no fixed times of discontinuity,*
- *the sequence of subdivisions $(\pi_n)_{n\ge1}$ in Assumption 2 is refining and each fixed time of discontinuity of X belongs to $\cap_n\pi_n$.*

Let M be a continuous $\mathbb{F}$ martingale such that $E(\sup_{s\le T}|M_s|)\le K$ and $E(\int_0^T|dA_s^{(n)}|)\le K$ for some K and all n, with A^n as in (1.8). Then

(i) M is a $\mathbb{G}^0$ special semimartingale.

(ii) Moreover, if $\mathbb{F}$ is the natural filtration of some càdlàg process Z, then M is a $\mathbb{G}$ special semimartingale with canonical decomposition $M=N+A$ such that N is a $\mathbb{G}$ martingale and $\sup_{0\le s\le T}|N_s|$ and $\int_0^T|dA_s|$ are integrable.

Proof. We construct the discretized process X^n defined by $X_t^n=X_{t_k^n}$ for all $t_k^n\le t<t_{k+1}^n$. That is

$$X_t^n=\sum_{i=0}^n X_{t_i^n}1_{\{t_i^n\le t<t_{i+1}^n\}}+X_T1_{\{t=T\}}$$

with the convention $t_0^n = 0$ and $t_{n+1}^n = T$. Let $\mathbb{G}^n$ be the smallest right-continuous filtration containing $\mathbb{F}$ and to which X^n is adapted.

Now, for $0 \le t \le T$,

$$\begin{aligned} X_t^n &= \sum_{i=0}^n X_{t_i^n} 1_{\{t_i^n \le t < t_{i+1}^n\}} + X_T 1_{\{t=T\}} = \sum_{i=0}^n X_{t_i^n} 1_{\{t_i^n \le t\}} - \sum_{i=0}^n X_{t_i^n} 1_{\{t_{i+1}^n \le t\}} + X_T 1_{\{t=T\}} \\ &= \sum_{i=1}^n (X_{t_i^n} - X_{t_{i-1}^n}) 1_{\{t_i^n \le t\}} + X_0 1_{\{t_0^n \le t\}} - X_{t_n^n} 1_{\{t_{n+1}^n \le t\}} + X_T 1_{\{t_{n+1}^n \le t\}} \\ &= X_0 1_{\{t_0^n \le t\}} + \sum_{i=1}^{n+1} (X_{t_i^n} - X_{t_{i-1}^n}) 1_{\{t_i^n \le t\}} = \sum_{i=0}^{n+1} (X_{t_i^n} - X_{t_{i-1}^n}) 1_{\{t_i^n \le t\}} \end{aligned}$$

with the notation $X_{t_{-1}^n} = 0$.

For each $0 \le i \le n+1$, let $\mathbb{H}^{i,n}$ be the initial expansion of $\mathbb{F}$ with $(X_{t_k^n} - X_{t_{k-1}^n})_{0 \le k \le i}$. Since $(X_{t_k^n} - X_{t_{k-1}^n})_{0 \le k \le i}$ satisfies Jacod's criterion, it follows that for each $0 \le i \le n+1$, $M - A^{i,n}$ is an $\mathbb{H}^{i,n}$ local martingale. Let

$$\tilde{\mathcal{G}}_t^n = \bigcap_{u > t} \mathcal{F}_u \vee \sigma\big((X_{t_i^n} - X_{t_{i-1}^n}) 1_{\{t_i^n \le u\}}, i = 0, \ldots, n+1\big)$$

Since the times t_k^n are fixed, $\mathbb{H}^{i,n}$ is also the initial expansion of $\mathbb{F}$ with $(t_k^n, X_{t_k^n} - X_{t_{k-1}^n})_{0 \le k \le i}$ and $\tilde{\mathbb{G}}^n = \mathbb{G}^n$ using a Monotone Class argument and the fact that $X_{t_k^n}^n = X_{t_k^n}$, for all $0 \le k \le n+1$. So it follows from Theorem 19 that $M - A^{(n)}$ is a $\mathbb{G}^n$ local martingale. An application of Theorem 25 yields the result. ■

Remark 3 *We refrain from stating Theorem 27 in a more general form for clarity but provide two extensions in the remarks below.*

(i) *Going beyond the continuous case for the $\mathbb{F}$ local martingale M is straightforward. We only need to use Theorem 19 in its general version rather than its application to the continuous case. However the explicit form of $A^{(n)}$ is much more complicated, which makes it hard to check the integrability assumption of Theorem 27. To be more concrete, one has to replace $A^{(n)}$ in the theorem above by $\tilde{A}^{(n)}$ defined by*

$$\tilde{A}_t^{(n)} = \sum_{i=0}^n \int_{t \wedge t_i^n}^{t \wedge t_{i+1}^n} (d\tilde{A}_s^{i,n} + dJ_s^{i,n})$$

i.e.

$$\tilde{A}_t^{(n)} = \sum_{i=0}^n 1_{\{t_i^n \le t < t_{i+1}^n\}} \Big(\sum_{k=0}^{i-1} \int_{t_k^n}^{t_{k+1}^n} (d\tilde{A}_s^{k,n} + dJ_s^{k,n}) + \int_{t_i^n}^{t} (d\tilde{A}_s^{i,n} + dJ_s^{i,n}) \Big)$$

where $\tilde{A}^{i,n}$ is the compensator of M in $\mathbb{H}^{i,n}$ as given by Jacod's theorem (see Theorems VI.10 and VI.11 in [117]) and $J^{i,n}$ is the dual predictable projection of $\Delta M_{t_{i+1}^n} 1_{[t_{i+1}^n, \infty[}$ onto $\mathbb{H}^{i,n}$.

(ii) A careful study of the proof above shows that Assumption 2 is only used to ensure that there exists an $\mathbb{H}^{i,n}$ predictable process $A^{i,n}$ such that $M - A^{i,n}$ is an $\mathbb{H}^{i,n}$ local martingale. Therefore, Theorem 27 will hold whenever this weaker assumption is satisfied.

If the sequence of filtrations $\mathbb{G}^n$ converges weakly to $\mathbb{G}$ then $(M - A^{(n)}, A^{(n)})$ converges in probability under the Skorohod J_1 topology to (N, A). Many criteria for this to hold are provided in the literature, see for instance[29, Propositions 3 and 4]. This holds for example when every $\mathbb{G}$ martingale is continuous and the subdivision $(\pi_n)_{n\geq 1}$ is refining. In this case, for each $0 \leq t \leq T$, $(\mathcal{G}_t^n)_{n\geq 1}$ is increasing and converges weakly to the σ-field $\mathcal{G}_t$. The following lemma allows us to conclude. See [29] for a proof.

Lemma 33 *Assume that every $\mathbb{G}$ martingale is continuous and that for every $0 \leq t \leq T$, $(\mathcal{G}_t^n)_{n\geq 1}$ increases (or decreases) and converges weakly to $\mathcal{G}_t$. Then $\mathbb{G}^n \xrightarrow{w} \mathbb{G}$.*

1.4.3 Examples: Time reversed diffusions and Kohatsu-Higa's example

1.4.3.1 Time reversed diffusions

Start with a Brownian filtration $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}$, $\mathcal{F}_t = \sigma(B_s, s \leq t)$ and consider the stochastic differential equation

$$dX_t = \sigma(X_t)dB_t + b(X_t)dt$$

Assume the existence of a unique strong solution $(X_t)_{0\leq t\leq T}$. Indeed, assume the transition density $\pi(t, x, y)$ exists and is twice continuously differentiable in x and continuous in t and y. This is guaranteed for example if b and σ are infinitely differentiable with bounded derivatives and if the Hörmander condition holds for any x (see [10]), and we assume that this holds in the sequel. In this case, π is even infinitely differentiable.

We next show how we can expand a filtration dynamically as t increases, via another stochastic process evolving backwards in time. To this end, define the time reversed process $Z_t = X_{T-t}$, for all $0 \leq t \leq T$. Let $\mathbb{G} = (\mathcal{G}_t)_{0\leq t<\frac{T}{2}}$ be the smallest right-continuous filtration containing $(\mathcal{F}_t)_{0\leq t<\frac{T}{2}}$ and to which $(Z_t)_{0\leq t<\frac{T}{2}}$ is adapted. We would like to prove that B remains a special semimartingale in $\mathbb{G}$ and give its canonical decomposition. That B is a $\mathbb{G}$ semimartingale can be obtained using the usual results from the filtration expansion theory. However, our approach allows us to obtain the decomposition, too. We assume (w.l.o.g) that $T = 1$. Introduce the reversed Brownian motion $\tilde{B}_t = B_{1-t} - B_1$ and the filtration $\tilde{\mathbb{G}} = (\tilde{\mathcal{G}}_t)_{0\leq t<\frac{1}{2}}$ defined by

$$\tilde{\mathcal{G}}_t = \bigcap_{t<u<\frac{1}{2}} \sigma(B_s, \tilde{B}_s, 0 \leq s < u).$$

Theorem 28 *Both B and $\tilde{B}$ are $\mathbb{G}$ semimartingales.*

Proof. First, it is well known that $\tilde{B}$ is a Brownian motion in its own natural filtration and $\sigma(B_{1-s} - B_1, 0 \leq s < \frac{1}{2})$ is independent from $\sigma(B_s, 0 \leq s < \frac{1}{2})$. Therefore $(B_t)_{0 \leq t < \frac{1}{2}}$ and $(\tilde{B}_t)_{0 \leq t < \frac{1}{2}}$ are independent Brownian motions in $\tilde{\mathbb{G}}$. Now, given our strong assumptions on the coefficients b and σ, X_1 satisfies Jacod's criterion with respect to $\tilde{\mathbb{G}}$. Therefore B and $\tilde{B}$ remain semimartingales in $\mathbb{H} = (\mathcal{H}_t)_{0 \leq t < \frac{1}{2}}$ where $\mathcal{H}_t = \bigcap_{\frac{1}{2} > u > t} \tilde{\mathcal{G}}_u \vee \sigma(X_1)$. It only remains to prove that $\mathbb{G} = \mathbb{H}$. For this, use [117, Theorem V.23] to get that

$$dX_{1-t} = \sigma(X_{1-t}) d\tilde{B}_t + (\sigma^{'}(X_{1-t})\sigma(X_{1-t}) + b(X_{1-t}))dt$$

Since $b + \sigma\sigma^{'}$ and σ are Lipschitz, $\bigcap_{\frac{1}{2} > u > t} \sigma(X_{1-s}, 0 \leq s \leq u) = \bigcap_{\frac{1}{2} > u > t} \sigma(\tilde{B}_s, 0 \leq s \leq u) \vee \sigma(X_1)$ and the result follows. ■

We apply now our results to obtain the $\mathbb{G}$ decomposition. This is the primary result of this section.

Theorem 29 *Assume there exists a nonnegative function ϕ such that $\int_0^1 \phi(s)ds < \infty$ and for each $0 \leq s < t$,*

$$E\Big(\Big|\frac{1}{\pi}\frac{\partial \pi}{\partial x}(t-s, X_s, X_t)\Big|\Big) \leq \phi(t-s)$$

Then the process $(B_t)_{0 \leq t < \frac{1}{2}}$ is a $\mathbb{G}$ semimartingale and

$$B_t - \int_0^t \frac{1}{\pi}\frac{\partial \pi}{\partial x}(1-2s, X_s, X_{1-s})ds$$

is a $\mathbb{G}$ Brownian motion.

Proof. Since the process Z_t is a càdlàg process with no fixed times of discontinuity, we can apply Theorem 27. First we prove that $(Z_t)_{0 \leq t < \frac{1}{2}}$ and $(\mathcal{F}_t)_{0 \leq t < \frac{1}{2}}$ satisfy Assumption 2. Let $(\pi_n)_{n \geq 1} = (\{t_i^n\})_{n \geq 1}$ be a refining sequence of subdivisions of $[0, \frac{1}{2}]$ whose mesh tends to zero. We will do more and compute directly the conditional distributions of $(Z_{t_0^n}, Z_{t_1^n} - Z_{t_0^n}, \ldots, Z_{t_i^n} - Z_{t_{i-1}^n})$ for any $1 \leq i \leq n+1$. Pick such i and let $0 \leq t < \frac{1}{2}$ and $(z_0, \ldots, z_i) \in \mathbb{R}^{i+1}$.

$$\begin{aligned}
&P(Z_{t_0^n} \leq z_0, Z_{t_1^n} - Z_{t_0^n} \leq z_1, \ldots, Z_{t_i^n} - Z_{t_{i-1}^n} \leq z_i \mid \mathcal{F}_t) \\
&\quad = P(X_1 \leq z_0, X_1 - X_{1-t_1^n} > -z_1, \ldots, X_{1-t_{i-1}^n} - X_{1-t_i^n} > -z_i \mid \mathcal{F}_t) \\
&\quad = E\Big(\prod_{k=1}^{i-1} 1_{\{X_{1-t_k^n} - X_{1-t_{k+1}^n} \geq z_{k+1}\}} P\big(X_{1-t_1^n} - z_1 \leq X_1 \leq z_0 \mid \mathcal{F}_{1-t_1^n}\big) \mid \mathcal{F}_t\Big) \\
&\quad = E\Big(\prod_{k=1}^{i-1} 1_{\{X_{1-t_k^n} - X_{1-t_{k+1}^n} \geq z_{k+1}\}} \int_{X_{1-t_1^n} - z_1}^{\infty} 1_{\{u_1 \leq z_0\}} P_{X_{1-t_1^n}}(t_1^n, u_1) du_1 \mid \mathcal{F}_t\Big) \\
&\quad = E\Big(\prod_{k=1}^{i-1} 1_{\{X_{1-t_k^n} - X_{1-t_{k+1}^n} \geq z_{k+1}\}} \int_{-z_1}^{\infty} 1_{\{v_1 \leq z_0 - X_{1-t_1^n}\}} P_{X_{1-t_1^n}}(t_1^n, v_1 + X_{1-t_1^n}) dv_1 \mid \mathcal{F}_t\Big)
\end{aligned}$$

Repeating the same technique and conditioning successively w.r.t $\mathcal{F}_{1-t_2^n}, \dots, \mathcal{F}_{1-t_i^n}$ gives

$$P(Z_{t_0^n} \le z_0, Z_{t_1^n} - Z_{t_0^n} \le z_1, \dots, Z_{t_i^n} - Z_{t_{i-1}^n} \le z_i \mid \mathcal{F}_t) = E\Big(\int_{-z_i}^{\infty} \cdots \int_{-z_1}^{\infty}$$
$$1_{\{\sum_{k=1}^i v_k \le z_0 - X_{1-t_i^n}\}} \prod_{k=1}^i P_{X_{1-t_i^n} + \sum_{j=k+1}^i v_j}(t_k^n - t_{k-1}^n, \sum_{l=k}^i v_l + X_{1-t_i^n}) dv_1 \dots dv_i \mid \mathcal{F}_t\Big)$$
$$= \int_{-\infty}^{\infty} \int_{-z_i}^{\infty} \cdots \int_{-z_1}^{\infty} 1_{\{u + \sum_{k=1}^i v_k \le z_0\}} P_{X_t}(1 - t_i^n - t, u)$$
$$\prod_{k=1}^i P_{u + \sum_{j=k+1}^i v_j}(t_k^n - t_{k-1}^n, u + \sum_{l=k}^i v_l) dv_1 \dots dv_i du$$

Fubini's Theorem implies then

$$P(Z_{t_0^n} \le z_0, Z_{t_1^n} - Z_{t_0^n} \le z_1, \dots, Z_{t_i^n} - Z_{t_{i-1}^n} \le z_i \mid \mathcal{F}_t) = \int_{-z_i}^{\infty} \cdots \int_{-z_1}^{\infty} \int_{-\infty}^{z_0 - \sum_{k=1}^i v_k}$$
$$P_{X_t}(1 - t_i^n - t, u) \prod_{k=1}^i P_{u + \sum_{j=k+1}^i v_j}(t_k^n - t_{k-1}^n, u + \sum_{l=k}^i v_l) du dv_1 \dots dv_i$$

Since the transition density $\pi(t, x, y) = P_x(t, y)$ is twice continuously differentiable in x by assumption, it is straightforward to check that

$$p_t^{i,n}(z_0, \dots, z_i) = \prod_{k=1}^i \pi(t_k^n - t_{k-1}^n, \sum_{j=0}^k z_j, \sum_{j=0}^{k-1} z_j) \pi(1 - t_i^n - t, X_t, \sum_{j=0}^i z_j)$$

One then readily obtains

$$d\langle p_\cdot^{i,n}(z_0, \dots, z_i), B_\cdot \rangle_s = \frac{1}{\pi} \frac{\partial \pi}{\partial x}(1 - t_i^n - s, X_s, \sum_{j=0}^i z_k) p_s^{i,n}(z_0, \dots, z_i) ds$$

Hence by taking the local martingale M in (1.7) to be B, we get

$$A_t^{i,n} = \int_0^t \frac{1}{\pi} \frac{\partial \pi}{\partial x}(1 - t_i^n - s, X_s, X_{1-t_i^n}) ds$$

Now equation (1.8) becomes

$$A_t^{(n)} = \sum_{i=0}^n 1_{\{t_i^n \le t < t_{i+1}^n\}} \Big(\sum_{k=0}^{i-1} \int_{t_k^n}^{t_{k+1}^n} \frac{1}{\pi} \frac{\partial \pi}{\partial x}(1 - t_k^n - s, X_s, X_{1-t_k^n}) ds$$
$$+ \int_{t_i^n}^t \frac{1}{\pi} \frac{\partial \pi}{\partial x}(1 - t_i^n - s, X_s, X_{1-t_i^n}) ds \Big)$$

In order to apply Theorem 27, it only remains to prove that $E(\int_0^{\frac{1}{2}} |dA_s^{(n)}|) \leq K$ for some constant K independent from n. The finite constant $K = \int_0^1 \phi(s)ds$ works since

$$\begin{aligned} E\Big(\int_0^{\frac{1}{2}} |dA_s^{(n)}|\Big) &\leq \sum_{i=0}^{n} \int_{t_k^n}^{t_{k+1}^n} E\Big|\frac{1}{\pi}\frac{\partial \pi}{\partial x}(1 - t_k^n - s, X_s, X_{1-t_k^n})\Big| ds \\ &\leq \sum_{i=0}^{n} \int_{t_k^n}^{t_{k+1}^n} \phi(1 - t_k^n - s) ds = \sum_{i=0}^{n} \int_{1-t_k^n-t_{k+1}^n}^{1-2t_k^n} \phi(s) ds \\ &\leq \sum_{i=0}^{n} \int_{1-2t_{k+1}^n}^{1-2t_k^n} \phi(s) ds = \int_0^1 \phi(s) ds \end{aligned}$$

This proves again that B is a $\mathbb{G}$ semimartingale. Now $A^{(n)}$ converges in probability to the process A given by

$$A_t = \int_0^t \frac{1}{\pi}\frac{\partial \pi}{\partial x}(1 - 2s, X_s, X_{1-s}) ds$$

Since all $\mathbb{G}$ martingales are continuous, the comment following Theorem 27 ensures that B-A is a $\mathbb{G}$ martingale. Its quadratic variation is t, therefore it is a $\mathbb{G}$ Brownian motion. ∎

In the Brownian case, the result in Theorem 29 can also be obtained using the usual theory of initial expansion of filtration. Assume $b = 0$ and $\sigma = 1$, i.e. $Z = B_{1-\cdot}$ and $X = B$.

Theorem 30 *The process B is a $\mathbb{G}$ semimartingale and*

$$B_t - \int_0^t \frac{B_{1-s} - B_s}{1 - 2s} ds, \qquad 0 \leq t < \frac{1}{2}$$

is a $\mathbb{G}$ Brownian motion.

Proof. Introduce the filtration $\mathbb{H}^1 = (\mathcal{H}_t)_{0 \leq t < \frac{1}{2}}$ obtained by initially expanding $\mathbb{F}$ with $B_{\frac{1}{2}}$.

$$\mathcal{H}_t^1 = \bigcap_{u>t} F_u \vee \sigma(B_{\frac{1}{2}})$$

We know that B remains an $\mathbb{H}^1$ semimartingale and

$$M_t := B_t - \int_0^t \frac{B_{\frac{1}{2}} - B_s}{\frac{1}{2} - s} ds, \qquad 0 \leq t < \frac{1}{2}$$

is an $\mathbb{H}^1$ Brownian motion. Now expand initially $\mathbb{H}^1$ with the independent σ-field $\sigma(B_v - B_{\frac{1}{2}}, \frac{1}{2} < v \leq 1)$ to obtain $\mathbb{H}$ i.e.

$$\mathcal{H}_t = \bigcap_{u>t} H_u^1 \vee \sigma(B_v - B_{\frac{1}{2}}, \frac{1}{2} < v \leq 1)$$

Obviously $(M_t)_{0\leq t<\frac{1}{2}}$ remains an $\mathbb{H}$ Brownian motion. But $\mathcal{G}_t \subset \mathcal{H}_t$, for all $0 \leq t < \frac{1}{2}$, hence the optional projection of M onto $\mathbb{G}$, denoted ${}^oM$ in the sequel, is again a martingale (see [50]), i.e.

$$^oM_t = B_t - E(\int_0^t \frac{B_{\frac{1}{2}} - B_s}{\frac{1}{2} - s} ds \mid \mathcal{G}_t), \qquad 0 \leq t < \frac{1}{2}$$

is a $\mathbb{G}$ martingale. Also, $N_t := E(\int_0^t \frac{B_{\frac{1}{2}}-B_s}{\frac{1}{2}-s} ds \mid \mathcal{G}_t) - \int_0^t E(\frac{B_{\frac{1}{2}}-B_s}{\frac{1}{2}-s} \mid \mathcal{G}_s) ds$ is a $\mathbb{G}$ local martingale, see for example [106] for a proof. So

$$B_t \;=\; {}^oM_t + N_t + \int_0^t E(\frac{B_{\frac{1}{2}} - B_s}{\frac{1}{2} - s} \mid \mathcal{G}_s) ds$$

We prove now the theorem using properties of the Brownian bridge. Recall that for any $0 \leq T_0 < T_1 < \infty$,

$$\mathcal{L}\Big((B_t)_{T_0\leq t\leq T_1} \mid B_s, s \notin]T_0, T_1[\Big) = \mathcal{L}\Big((B_t)_{T_0\leq t\leq T_1} \mid B_{T_0}, B_{T_1}\Big) \tag{1.9}$$

and

$$\mathcal{L}\Big((B_t)_{T_0\leq t\leq T_1} \mid B_{T_0} = x, B_{T_1} = y\Big) = \mathcal{L}\Big(x + \frac{t - T_0}{T_1 - T_0}(y - x) + (Y^{W,T_1-T_0}_{t-T_0})_{T_0\leq t\leq T_1}\Big)$$

where W is a generic standard Brownian motion and Y^{W,T_1-T_0} is the standard Brownian bridge on $[0, T_1 - T_0]$. It follows that for all $T_0 \leq t \leq T_1$ and all x and y,

$$E(B_t \mid B_{T_0} = x, B_{T_1} = y) = \frac{T_1 - t}{T_1 - T_0} x + \frac{t - T_0}{T_1 - T_0} y \tag{1.10}$$

For any $0 \leq s < t < \frac{1}{2}$, if follows from (1.9) and (1.10) that

$$E(B_{\frac{1}{2}} - B_s \mid \mathcal{G}_s) = \frac{1}{2}(B_{1-s} - B_s)$$

Therefore

$$B_t - \int_0^t \frac{B_{1-s} - B_s}{1 - 2s} ds \;=\; {}^oM_t + N_t$$

is a $\mathbb{G}$ local martingale. Since the quadratic variation of the $\mathbb{G}$ local martingale $B - A$ is t, Levy's characterization of Brownian motion ends the proof. ∎

In the immediately previous proof, the properties of the Brownian bridge allow us to compute explicitly the decomposition of B in $\mathbb{G}$. Our method obtains both the semimartingale property and the decomposition simultaneously and generalizes to diffusions, for which the computations as in the proof of Theorem 30 are hard. We provide a shorter proof for Theorem 30 based on Theorem 29. This illustrated that, given Theorem 29, even in the Brownian case our method is shorter, simpler, and more intuitive.

Proof. [Second proof of Theorem 30] In the Brownian case, $\pi(t,x,y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(y-x)^2}{2t}}$. Therefore $\frac{1}{\pi}\frac{\partial \pi}{\partial x}(t,x,y) = \frac{y-x}{t}$. Hence

$$E\Big(|\frac{1}{\pi}\frac{\partial \pi}{\partial x}|(t-s, B_s, B_t)\Big) \leq \frac{1}{t-s} E(|B_t - B_s|) = \sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{t-s}}$$

and $\phi(x) = \sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{x}}$ is integrable in zero. From the closed formula for the transition density, $A_t = \int_0^t \frac{B_{1-s}-B_s}{1-2s} ds$. Therefore B is a $\mathbb{G}$ semimartingale, and $B-A$ is a $\mathbb{G}$ Brownian motion by Theorem 29. ■

This property satisfied by Brownian motion is inherited by diffusions whose parameters b and σ satisfy some boundedness assumptions. We add the extra assumptions that b and σ are bounded and $k \leq \sigma(x)$ for some $k > 0$. The following holds.

Corollary 13 *The process $(B_t)_{0\leq t<\frac{1}{2}}$ is a $\mathbb{G}$ semimartingale and*

$$B_t - \int_0^t \frac{1}{\pi}\frac{\partial \pi}{\partial x}(1-2s, X_s, X_{1-s}) ds$$

is a $\mathbb{G}$ Brownian motion.

Proof. Introduce the following quantities

$$s(x) = \int_0^x \frac{1}{\sigma(y)} dy \qquad g = s^{-1} \qquad \mu = \frac{b}{\sigma}\circ g - \frac{1}{2}\sigma' \circ g$$

The process $Y_t = s(X_t)$ satisfies the SDE $dY_t = \mu(Y_t)dt + dB_t$. The transition density is known in semi-closed form (see [49]) and given by

$$\pi(t,x,y) = \frac{1}{\sqrt{2\pi t}}\frac{1}{\sigma(y)} e^{-\frac{(s(y)-s(x))^2}{2t}} U_t(s(x), s(y))$$

where $U_t(x,y) = H_t(x,y)e^{A(y)-A(x)}$, $H_t(x,y) = E(e^{-t\int_0^1 h(x+z(y-x)+\sqrt{t}W_z)dz})$, W is a Brownian bridge, A a primitive of μ and $h = \frac{1}{2}(\mu^2 + (\mu')^2)$. It is then straightforward to compute the ratio

$$\begin{aligned}
\frac{1}{\pi}\frac{\partial \pi}{\partial x}(t,x,y) &= \frac{1}{\sigma(x)}\Big(\frac{s(y)-s(x)}{t} + \frac{1}{U_t(s(x),s(y))}\frac{\partial U_t}{\partial x}(s(x),s(y))\Big) \\
&= \frac{1}{\sigma(x)}\Big(\frac{s(y)-s(x)}{t} + \frac{1}{H_t(s(x),s(y))}\frac{\partial H_t}{\partial x}(s(x),s(y)) - \mu(s(x))\Big)
\end{aligned}$$

From the boundedness assumptions of b and σ and their derivatives, there exists a constant M such that $|\frac{1}{\pi}\frac{\partial \pi}{\partial x}(t,x,y)| \leq M(1 + \frac{|s(y)-s(x)|}{t})$. Hence, for $0 \leq s < t$

$$E\Big|\frac{1}{\pi}\frac{\partial \pi}{\partial x}(t-s, X_s, X_t)\Big| \leq M\big(1 + E\Big|\frac{s(X_t)-s(X_s)}{t-s}\Big|\big)$$

But $s(X_t) - s(X_s) = Y_t - Y_s = \int_s^t \mu(Y_u)du + W_t - W_s$. But μ is bounded, hence

$$E|s(Y_t) - s(Y_s)| \leq ||\mu||_\infty |t-s| + E|W_t - W_s| = ||\mu||_\infty |t-s| + \sqrt{\frac{2}{\pi}}\sqrt{t-s}$$

This proves the existence of a constant C such that

$$E\Big|\frac{1}{\pi}\frac{\partial \pi}{\partial x}(t-s, X_s, X_t)\Big| \leq C(1 + \frac{1}{\sqrt{t-s}})$$

Since $\phi(x) = C(1 + \frac{1}{\sqrt{x}})$ is integrable in zero, we can apply Theorem 29 and conclude. ■

1.4.3.2 Stochastic volatility models

Let $(\Omega, \mathcal{H}, P, \mathbb{H})$ be a filtered probability space. Assume that we are given an $\mathbb{H}$ Brownian motion W and a positive continuous $\mathbb{H}$ adapted process σ. Consider the following stochastic volatility model

$$dS_t = S_t \sigma_t dW_t$$

and σ is such that (σ, A) is Markov w.r.t its natural filtration with transition density $p_t((u,a),(v,b))$, where $A_t := \int_0^t \sigma_s^2 ds$. Define $Z_t = \int_0^t \sigma_s dW_s$, so that $S_t = \mathcal{E}(Z)_t$ and let $\mathbb{F}$ be the filtration generated by S and σ. Then $\mathcal{F}_t = \bigcap_{u>t} \sigma(Z_s, \sigma_s, s \leq u)$. Since $\langle Z, Z\rangle_t = A_t$, A is $\mathbb{F}$ adapted.

We want to expand $(\mathcal{F}_t)_{0 \leq t < \frac{T}{2}}$ progressively with the continuous process $X_t := \int_{T-t}^T \sigma_s^2 ds = A_T - A_t$, $0 \leq t < \frac{1}{2}$. The process $(X_t)_{0 \leq t < \frac{T}{2}}$ satisfies Assumption 2 with density

$$\int_{a_1=0}^\infty \int_{u_1=0}^\infty \cdots \int_{u_{n+1}=0}^\infty \prod_{k=1}^n p_{t_{k+1}^n - t_k^n}\big((u_k, A_t + \sum_{i=1}^k a_i), (u_{k+1}, A_t + \sum_{i=1}^{k+1} a_i)\big)$$
$$p_{\frac{T}{2}-t}\big((\sigma_t, A_t), u_1, a_1 + A_t)\big) du_{n+1} \dots du_1 da_1$$

along any refining subdivision $\{t_k^n\}$ of $[0, \frac{T}{2}]$. Let M be a continuous $\mathbb{F}$ martingale such that $\sup_s |M_s|$ is integrable. Then $A^{i,n}$ and $A^{(n)}$ can be computed using the formulas in (1.7) and (1.8) and the $\mathbb{G}$ semimartingale property of M will be guaranteed as soon as $E(\int_0^{\frac{T}{2}} |dA_s^{(n)}|) \leq K$ for some K and all n.

1.4.3.3 Time reversed diffusions : an extension

We studied above in detail the expansion of $\mathbb{F}$ with a time reversed diffusion $X_t = Z_{T-t}$, up to time $\frac{T}{2}$. A first extension consists in making time move backward at a different speed than it moves forward, that is, let A be an increasing process and expand $\mathbb{F}$ with

$X_t = Z_{A(T-t)}$ where Z is the solution to some SDE driven by W to get $\mathbb{G}$. Assume Z is the unique strong solution to

$$dZ_t = \sigma(Z_t)dW_t + b(Z_t)dt$$

Assume the transition density $\pi(t,x,y)$ exists and is twice continuously differentiable in x and continous in both t and y. Assume A is an increasing positive continuous function such that $A(0) < T$, so that there exists a unique $0 < t_0 < T$ such that $t_0 = A(T - t_0)$. Then W remains a $\mathbb{G}$ semimartingale up to time t_0. The following theorem makes this more precise.

Theorem 31 *Assume there exists a nonnegative function ϕ such that $\int_0^{A(T)} \phi(s)ds < \infty$ and for $s < t$,*

$$E\Big(|\frac{1}{\pi}\frac{\partial \pi}{\partial x}(t-s, Z_s, Z_t|\Big) \leq \phi(t-s)$$

Then the process $(W_t)_{0\leq t<t_0}$ is a $\mathbb{G}$ semimartingale and

$$W_t - \int_0^t \frac{1}{\pi}\frac{\partial \pi}{\partial x}(A(T-s) - s, Z_s, X_s)ds$$

is a $\mathbb{G}$ Brownian motion.

Proof. The proof follows the same lines as the proof of Theorem 29 and is therefore sketched briefly. Let $(\pi_n)_{n\geq 1} = (\{t_i^n\})_{n\geq 1}$ be a refining sequence of subdivisions of $[0, t_0]$ whose mesh tends to zero. The process $A^{(n)}$ in (1.8) is given by

$$A_t^{(n)} = \sum_{i=0}^{n} 1_{\{t_i^n \leq t < t_{i+1}^n\}}\Big(\sum_{k=0}^{i-1}\int_{t_k^n}^{t_{k+1}^n} \frac{1}{\pi}\frac{\partial \pi}{\partial x}(A(T-t_k^n) - s, Z_s, X_{t_k^n})ds$$
$$+ \int_{t_i^n}^{t} \frac{1}{\pi}\frac{\partial \pi}{\partial x}(A(T-t_i^n) - s, Z_s, X_{t_i^n})ds\Big)$$

and $E(\int_0^{t_0} |dA_s^{(n)}|) \leq \sum_{k=0}^n \int_{t_k^n}^{t_{k+1}^n} \phi(A(T-t_k^n) - s)ds$. Changing variables and using that A is nondecreasing, it follows that

$$E(\int_0^{t_0} |dA_s^{(n)}|) \leq \sum_{k=0}^{n} \int_{A(T-t_{k+1}^n)-t_{k+1}^n}^{A(T-t_k^n)-t_k^n} \phi(u)du = \int_0^{A(T)} \phi(u)du$$

since $A(T - t_0) = t_0$. Therefore W is a $\mathbb{G}$ semimartingale. Since $A^{(n)}$ converges in probability to the process $A_t := \int_0^t \frac{1}{\pi}\frac{\partial\pi}{\partial x}(A(T-s)-s, Z_s, X_s)ds$ and since all $\mathbb{G}$ martingales are continuous, the claim follows from Theorem 27. ■

Exactly as in Corollary 13, the conditions of Theorem 31 can be proved to hold for $Z = W$ (case $b = 0$ and $\sigma = 1$) and then proved to be satisfied more generally if the coefficients b and σ of the SDE are bounded and such that $k \leq \sigma(x)$ for some $k > 0$. A second extension of the filtration expansion with a time reversed diffusion idea would be to make the time change A stochastic. We refrain from introducing explicit examples because this extension would be only computationally more involved.

To provide a concrete application for insider trading models, fix $0 < \varepsilon < T$ and define $A(t) = T - \frac{T-t}{T-\varepsilon}\varepsilon$. In this case, $t_0 = T - \varepsilon$ and the trader acquires the information $(Z_t)_{T-\varepsilon<t\leq T}$ progressively (and backward in time) when time runs from 0 to $T-\varepsilon$. That is, the insider does not only hold the information Z_T as in classical models, but his information gets more precise as he discovers progressively the entire path of Z in a small interval close to maturity.

1.4.3.4 Kohatsu-Higa's example and extension

In [98], the author suggests an insider model, closer to reality than the classical insider models, where the information the insider holds gets deformed through time. More precisely, Kohatsu-Higa expands progressively the natural filtration $\mathbb{F}$ of a Brownian motion W with the process $X_t = W_T + Z_{(T-t)^\theta}$, where Z is a Brownian motion independent of W. In this case the blurring of the information is done at a logarithmic scale. The decomposition of W in the expanded filtration is provided and the proof relies on properties of Gaussian random vectors. For the sake of clarity, we focus on the same class of examples and show how the decomposition of W in the expanded filtration comes out clearly from our approach. Let A be an increasing continuous function, $X_t = W_T + Z_{A(T-t)}$ and let $\mathbb{G}$ be the progressive expansion of $\mathbb{F}$ with X. The following holds.

Theorem 32 (Kohatsu-Higa) *Assume $\int_0^T \frac{ds}{\sqrt{A(s)}} < \infty$. Then W is a $\mathbb{G}$ semimartingale and*

$$W_t - \int_0^t \frac{X_s - W_s}{T - s + A(T-s)} ds$$

is a $\mathbb{G}$ Brownian motion.

Proof. Let $(\pi_n)_{n\geq 1} = (\{t_i^n\})_{n\geq 1}$ be a refining sequence of subdivisions of $[0,T]$ whose mesh tends to zero. Let $1 \leq i \leq n+1$, and $0 \leq t < T$ and $(x_0, \dots, x_i) \in \mathbb{R}^{i+1}$. The $\mathcal{F}_t$ conditional density of $(X_{t_0^n}, X_{t_1^n} - X_{t_0^n}, \dots, X_{t_i^n} - X_{t_{i-1}^n})$ is given by

$$p_t(x_0, \dots, x_i) = \prod_{k=1}^{i} g(-x_k, A(T-t_k^n) - A(T-t_{k-1}^n)) \int_{-\infty}^{\infty} g(v, A(T-t_i^n)) g(\sum_{k=0}^{i} -W_t - v, T-t) dv$$

where g is the gaussian density. By completing the squares, we can compute the integral above, since more generally

$$\int_{-\infty}^{\infty} g(x+a, s) g(x+b, t) dx = g(b-a, s+t)$$

for all $(a, b) \in \mathbb{R}^2$ and $s, t > 0$. Therefore,

$$p_t(x_0, \dots, x_i) = \prod_{k=1}^{i} g(-x_k, A(T-t_k^n) - A(T-t_{k-1}^n)) g(W_t - \sum_{k=0}^{i} x_k, T - t + A(T-t_i^n))$$

Using Ito's formula, one gets

$$\frac{d\langle p(x_0,\dots,x_i),W\rangle_t}{p_t(x_0,\dots,x_i)} = \frac{\sum_{k=0}^{i} x_k - W_t}{T-t+A(T-t_i^n)}$$

Equation (1.8) becomes

$$A_t^{(n)} = \sum_{i=0}^{n} 1_{\{t_i^n \le t < t_{i+1}^n\}} \Big(\sum_{k=0}^{i-1} \int_{t_k^n}^{t_{k+1}^n} \frac{X_{t_k^n} - W_s}{T-s+A(T-t_k^n)} ds + \int_{t_i^n}^{t} \frac{X_{t_i^n} - W_s}{T-s+A(T-t_i^n)} ds \Big)$$

We prove now that the sequence $\big(E(\int_0^T |dA_s^{(n)}|)\big)_{n\ge 1}$ is uniformly bounded.

$$\begin{aligned} E(\int_0^T |dA_s^{(n)}|) &\le \sum_{i=0}^{n} \int_{t_i^n}^{t_{i+1}^n} E\Big|\frac{X_{t_i^n} - W_s}{T-s+A(T-t_i^n)}\Big| ds \\ &\le \sum_{i=0}^{n} \int_{t_i^n}^{t_{i+1}^n} \Big(\frac{E|W_T - W_s|}{T-s} + \frac{E|Z_{A(T-t_i^n)}|}{T-s+A(T-t_i^n)}\Big) ds \le \sqrt{\frac{2}{\pi}} \int_0^T \Big(\frac{ds}{\sqrt{T-s}} + \frac{ds}{\sqrt{A(T-s)}}\Big) \end{aligned}$$

which is finite by assumption. Therefore W is a semimartingale, and since all $\mathbb{G}$ martingales are continuous, it follows that $W - A$ is a martingale, where A is the limit in probability of $A^{(n)}$ and is therefore given by $A_t = \int_0^t \frac{X_s - W_s}{T-s+A(T-s)} ds$. That $W - A$ is a $\mathbb{G}$ Brownian motion follows from the fact that its quadratic variation is t. ■

We recover the example in [98] with the choice $A(t) = t^\theta$, at least for $\theta < 2$. However, the result remains valid for more general processes X, of the form $X_t = Y_T + Z_{A(T-t)}$, where $Y_T = f((W_s)_{0\le s\le T})$ is an $\mathcal{F}_T$ measurable random variable with an $\mathbb{F}$ conditional density g, Z is a process independent from $\mathbb{F}$ and A is an increasing function. With X having this structure, computations can usually be done explicitly and Assumption 2 and the integrability conditions of Theorem 27 can be checked. This provides a class of non trivial dynamic insider trading models which allow a general modeling of the deformation of the information through time.

1.5 Toward dynamic models for insider trading

Many processes will be such that Assumption 2 is not satisfied. In this section, we focus our attention on the case where the process X used to enlarge the base filtration $\mathbb{F}$ is continuous. In this case, new criterions that are sometimes easier to check can be found to ensure that a given $\mathbb{F}$ semimartingale M, satisfying some integrability assumptions, remains a semimartingale in the expanded filtration. Now, instead of assuming that the increments of X satisfy Assumption 2, we will instead work with the successive hitting times of X of some given levels. In case these times are either honest or initial, we can reach the same conclusions as in the previous section. We recover Jeulin's example, where the natural filtration of a Bessel 3 process is progressively expanded with its remaining infimum (see [89]) and we also extend this result to the case where we expand the natural filtration of a transient diffusion R with its remaining infimum.

We will use the examples of this section to point out how *riskless* arbitrage opportunities appear in this framework of filtration expansions. Take a continuous semimartingale $X_t = X_0 + M_t + A_t, 0 \leq t \leq T$ on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ where $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$. Here M is the continuous local martingale part and A is a process with paths of finite variation on compact time sets, almost surely. NFLVR can fail to hold mainly for two reasons and this is made rigourous in the following theorem, which provides necessary and sufficient conditions such that there exists an equivalent probability measure Q such that X is a Q local martingale. The proof can be found in [118].

Theorem 33 (Protter and Shimbo) *Let $X_t = X_0 + M_t + A_t, t \geq 0$ be a continuous semimartingale on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ and denote by C its quadratic variation. There exists an equivalent probability measure Q on $\mathcal{F}_T$ such that X is a Q local martingale only if*

(i) dA is absolutely continuous w.r.t dC

(ii) If J is such that $A_t = \int_0^t J_s dC_s$ for each $0 \leq t \leq T$, then $\int_0^T J_s^2 dC_s < \infty$ a.s.

If in addition one has the condition $E(\mathcal{E}(-\int_0^T J_s dM_s)) = 1$, then we have sufficient conditions for there to exist an equivalent probability measure Q on $\mathcal{F}_T$ such that X is a Q local martingale.

There are many examples in the literature of insider models that introduce arbitrage opportunities in the sense of FLVR by violating Condition (ii) of Theorem 33. See for instance [71] where a Brownian filtration is expanded with the supremum of the Brownian motion over a finite time horizon. The Brownian motion remains a semimartingale in the expanded filtration with a finite variation part that is absolutely continuous w.r.t Lebesgue measure but lacks enough of nice integrability properties that the insider information does introduce free lunches. In our example, we will notice that the finite variation part in the progressively expanded filtration becomes singular w.r.t Lebesgue measure and this provides a dynamic insider trading model with *extreme* arbitrage opportunities, from which the risk is removed and arising because Condition (i) of Theorem 33 is violated.

1.5.1 Filtration expansion results based on a Jacod's criterion for hitting times

Let $\mathbb{F}$ be the natural filtration of some càdlàg process and let X be a continuous process which is not adapted to $\mathbb{F}$. Let $\mathbb{G}$ be the progressive expansion of $\mathbb{F}$ with X. We provide in this subsection a Jacod's type criterion (for successive hitting times by X of some levels) that guarantees that a given $\mathbb{F}$ semimartingale M, satisfying some integrability assumptions, remains a $\mathbb{G}$ semimartingale. Some intermediate filtrations are needed. First, to take care of the initial value of the process X, we expand $\mathbb{F}$ initially with X_0 to obtain

$$\mathcal{F}_t^{X_0} = \bigcap_{u>t} \mathcal{F}_u \vee \sigma(X_0)$$

We need at least that M remains an $\mathbb{F}^{X_0}$ semimartingale. A sufficient condition that ensures this is that X_0 satisfies Jacod's criterion w.r.t $\mathbb{F}$. Throughout this subsection, we assume that the following holds.

Assumption 3 *X_0 satisfies Jacod's criterion w.r.t $\mathbb{F}$ i.e. the regular $\mathcal{F}_t$-conditional probabilities $P(X_0 \in dx \mid \mathcal{F}_t)$ are absolutely continuous w.r.t the law of X_0, for all $0 \le t \le T$.*

Now, since X_0 is $\mathcal{F}_t^{X_0}$ measurable, for each $0 \le t \le T$, we just need to expand $\mathbb{F}^{X_0}$ with $X - X_0$ to obtain $\mathbb{G}$. So as long as Assumption 3 is satisfied and up to replacing in the sequel $\mathbb{F}$ with $\mathbb{F}^{X_0}$, we can assume that the process X starts at 0 and we do so.

Now we approximate X with a counting process in the following way. Let $(\varepsilon_n)_{n\ge1}$ be a sequence of positive real numbers decreasing to zero. Define $\tau_0^n = 0$ and define τ_p^n recursively by

$$\tau_p^n = \inf\{t > \tau_{p-1}^n,\ |X_t - X_{\tau_{p-1}^n}| \ge \varepsilon_n\} \qquad \text{for each } p \ge 1$$

For each $n \ge 1$, $(\tau_p^n)_{p\ge1}$ is a strictly increasing sequence of random times. Define now the sequence of processes

$$X_t^n = \sum_{p=0}^{\infty} X_{\tau_p^n} 1_{\{\tau_p^n \le t < \tau_{p+1}^n\}}$$

for each $n \ge 1$. We need some assumptions on the random times τ_p^n, mainly that they do not cluster.

Assumption 4 *For each $n \ge 1$, the sequence $(\tau_p^n)_{p\ge1}$ is strictly increasing to infinity. That is, $\tau_p^n > \tau_{p-1}^n$ on the set where $\tau_{p-1}^n < \infty$ and $\lim_{p\to\infty} \tau_p^n = \infty$.*

Define the intermediate filtrations $\mathbb{G}^n = (\mathcal{G}_t^n)_{0\le t\le T}$ where

$$\mathcal{G}_t^n = \bigcap_{u>t} \mathcal{F}_u \vee \sigma(X_s^n, s \le u)$$

The following holds.

Lemma 34 *If Assumption 4 is satisfied, X^n converges a.s. to X and for each $0 \le t \le T$, $\mathcal{G}_t^n \xrightarrow{w} \mathcal{G}_t$.*

Proof. Since $\lim_{p\to\infty} \tau_p^n = \infty$, it follows that $|X_t^n - X_t| \le \sum_{p=0}^{\infty} |X_{\tau_p^n} - X_t| 1_{\{\tau_p^n \le t < \tau_{p+1}^n\}}$. Since on $\{\tau_p^n \le t < \tau_{p+1}^n\}$, $|X_t - X_{\tau_p^n}| \le \varepsilon_n$, it follows that $\sup_{0\le t\le T} |X_t^n - X_t| \le \varepsilon_n$ which yields the first claim. Now, the second claim follows as in Lemma 16. ∎

We assume in the sequel that Assumption 4 is satisfied. Theorem 27 can be used together with Lemma 34 above to prove that M remains a $\mathbb{G}$ special semimartingale as soon as one can guarantee that M remains, for each n, a $\mathbb{G}^n$ special semimartingale whose decomposition $M = M^n + A^n$ satisfies

$$E(\int_0^T |dA_s^n|) \leq K \qquad \text{and} \qquad E(\sup_{0\leq s\leq T} |M_s^n|) \leq K$$

for all $n \geq 1$ and some K independent from n. We assume in the sequel that $\sup_{0\leq s\leq T} |M_s|$ is integrable. Under this integrability assumption, we only need to find sufficient conditions on X such that the following condition is satisfied.

Condition 3 *The processes X and M are such that*

(i) For each $n \geq 1$, M is a $\mathbb{G}^n$ special semimartingale.

(ii) The finite variation term A^n in the $\mathbb{G}^n$ canonical decomposition of M satisfies $E(\int_0^T |dA_s^n|) \leq K$ for some K and all $n \geq 1$.

1.5.1.1 Condition 3 (i)

To deal with the $\mathbb{G}^n$ semimartingale property of M, i.e. Condition 3 (i), we approximate the processes X^n and introduce new intermediate filtrations. For each $n \geq 1$, consider the sequence $(X^{n,k})_{k\geq 1}$ of processes defined by

$$X_t^{n,k} = \sum_{p=1}^{k} \varepsilon_{n,p} 1_{\{\tau_p^n \leq t\}}$$

where $\varepsilon_{n,p} = \varepsilon_n \big(2 \cdot 1_{\{X_{\tau_p^n} > X_{\tau_{p-1}^n}\}} - 1\big)$, for each $p \geq 1$.

Lemma 35 *For each $n \geq 1$, $X^{n,k}$ converges a.s. to X^n.*

Proof.

$$\begin{aligned} X_t^n &= \sum_{p=0}^{\infty} X_{\tau_p^n} 1_{\{\tau_p^n \leq t < \tau_{p+1}^n\}} = \sum_{p=0}^{\infty} X_{\tau_p^n} 1_{\{\tau_p^n \leq t\}} - \sum_{p=1}^{\infty} X_{\tau_{p-1}^n} 1_{\{\tau_p^n \leq\}} \\ &= X_0 + \sum_{p=1}^{\infty} (X_{\tau_p^n} - X_{\tau_{p-1}^n}) 1_{\{\tau_p^n \leq t\}} = \sum_{p=1}^{\infty} (X_{\tau_p^n} - X_{\tau_{p-1}^n}) 1_{\{\tau_p^n \leq t\}} \end{aligned}$$

since $X_0 = 0$ by assumption. Now, since X is continuous,

$$X_{\tau_p^n} - X_{\tau_{p-1}^n} = \varepsilon_n 1_{\{X_{\tau_p^n} > X_{\tau_{p-1}^n}\}} - \varepsilon_n 1_{\{X_{\tau_p^n} \leq X_{\tau_{p-1}^n}\}} = \varepsilon_{n,p}$$

Therefore $X_t^n = \sum_{p=1}^{\infty} \varepsilon_{n,p} 1_{\{\tau_p^n \leq t\}}$. It is now clear that $X^{n,k}$ converges a.s. to X^n. ■

Introduce the filtrations $\mathbb{G}^{n,k} = (\mathcal{G}_t^{n,k})_{0\leq t\leq T}$ where

$$\mathcal{G}_t^{n,k} = \bigcap_{u>t} \mathcal{F}_u \vee \sigma(X_s^{n,k}, s \leq u)$$

First, notice that $\mathcal{G}_t^{n,k}$ can be represented in the following way.

Lemma 36 *For each $n \geq 1$ and $k \geq 1$,*

$$\mathcal{G}_t^{n,k} = \bigcap_{u>t} \mathcal{F}_u^{X_0} \vee \sigma(\tau_i^n \wedge s, s \leq u, 1 \leq i \leq k) \vee \sigma(\varepsilon_{n,i} 1_{\{\tau_i^n \leq s\}}, s \leq u, 1 \leq i \leq k)$$

For each $n \geq 1$, the sequence of filtrations $(\mathbb{G}^{n,k})_{k \geq 1}$ is increasing.

Lemma 37 *For each $n \geq 1$ and each $0 \leq t \leq T$, $\mathcal{G}_t^{n,k} \xrightarrow{w} \mathcal{G}_t^n$.*

Proof. This follows from Lemma 35 above using the same proof as in Lemma 16. Alternatively one can notice that $\bigvee_{k=1}^{\infty} \mathcal{G}_t^{n,k} = \mathcal{G}_t^n$ and it follows from Lemma 36 that the sequence of σ-fields $(\mathcal{G}_t^{n,k})_{k \geq 1}$ is increasing. Let $0 \leq t \leq T$ and let Z be an integrable $\mathcal{G}_t^n$ measurable random variable. Then $M_k = E(Z \mid \mathcal{G}_t^{n,k})$ is a closed martingale and the convergence theorem for closed martingales ensures that M_k converges to Z in L^1, which implies that $E(Z \mid \mathcal{G}_t^{n,k}) \xrightarrow{P} Z$. This means that $\mathcal{G}_t^{n,k} \xrightarrow{w} \mathcal{G}_t^n$. ∎

One then needs to guarantee that the following condition holds for each $n \geq 1$ to be able to use Theorem 27.

Condition 4 *The processes X and M are such that*

(i) For each $k \geq 1$, M is a $\mathbb{G}^{n,k}$ special semimartingale.

(ii) The finite variation terms $A^{n,k}$ in the $\mathbb{G}^{n,k}$ canonical decompositions of M satisfy $E(\int_0^T |dA_s^{n,k}|) \leq K$ for all $k \geq 1$ and some K independent from k.

If Condition 4 is satisfied then Condition 3 (i) is satisfied by Theorem 27. However, since the times $(\tau_p^n)_{p \geq 1}$ are strictly increasing to infinity, it is sufficient to guarantee that Condition 4 (i) only is satisfied. That is, the following holds.

Theorem 34 *Under Assumption 4, if Condition 4 (i) is satisfied for each $n \geq 1$, then Condition 3 (i) is satisfied.*

To prove this theorem, we need the following lemma which follows trivially from the results in section 1.2.3.4. We summarize the proof for clarity reasons.

Lemma 38 *The following holds.*

(i) For each k, n and t,

$$\{t < \tau_k^n\} \cap \mathcal{G}_t^{n,k-1} = \{t < \tau_k^n\} \cap \mathcal{G}_t^{n,k}$$

(ii) For $j \geq i \geq 1$, and for each t,

$$\{t < \tau_i^n\} \cap \mathcal{G}_t^{n,i-1} = \{t < \tau_i^n\} \cap \mathcal{G}_t^{n,j}$$

(iii) As a consequence,

$$\{t < \tau_i^n\} \cap \mathcal{G}_t^{n,i-1} = \{t < \tau_i^n\} \cap \mathcal{G}_t^n$$

(iv) *Equivalently, let $t > 0$ and $k \geq 1$be fixed and let h be a $\mathcal{G}_t^n$ measurable random variable. Then there exists a $\mathcal{G}_t^{n,k-1}$ measurable random variable f such that*

$$h1_{\{\tau_k^n > t\}} = f1_{\{\tau_k^n > t\}}$$

Proof. Since in all assertions, the index n is fixed, we can drop it. That is, consider a base filtration $\mathbb{F}$, a sequence of increasing times τ_k and a sequence of random variables X_k. Define a sequence of progressive expansions $\mathbb{G}^k$ through

$$\mathcal{G}_t^k = \mathcal{G}_t^{k-1,\tau_k} \vee \sigma(X_k 1_{\{\tau_k \leq s\}}, s \leq t)$$

where $\mathbb{G}^{k-1,\tau_k}$ is the progressive expansion of $\mathbb{G}^{k-1}$with τ_k and define $\mathbb{G}$ to be the filtration generated by all $\mathbb{G}^k$. It well known that $\{\tau_k > t\} \cap \mathcal{G}_t^{k-1,\tau_k} = \{\tau_k > t\} \cap \mathcal{G}_t^{k-1}$. Let $H_t = Y_t h(1_{\{\tau_k \leq t\}} X_k)$ with Y_t bounded and $\mathcal{G}_t^{k-1,\tau_k}$ measurable and h a bounded Borel function. Then $H_t 1_{\{\tau_k > t\}} = Y_t h(0) 1_{\{\tau_k > t\}}$ which is measurable w.r.t $\{\tau_k > t\} \cap \mathcal{G}_t^{k-1,\tau_k} = \{\tau_k > t\} \cap \mathcal{G}_t^{k-1}$. The Monotone Class Theorem now proves that $\{\tau_k > t\} \cap \mathcal{G}_t^k \subset \{\tau_k > t\} \cap \mathcal{G}_t^{k-1}$. The reverse inclusion is clear and (i) is proved. To prove (ii), we intersect both sides of (i) written for the index i, by $\{t < \tau_i\}$ and use that $\tau_i < \tau_{j+1}$ implies $\{t < \tau_i\} \subset \{t < \tau_{j+1}\}$ to get $\{t < \tau_i\} \cap \mathcal{G}_t^i = \{t < \tau_i\} \cap \mathcal{F}_t^{j+1}$. Hence if the statement is true for j, it is also true for $j + 1$. It is true for $j = i$ by (i), so the result follows by induction. Regarding (iii), since $\mathcal{G}_t$ is generated by all the $\mathcal{G}_t^i$, the result follows from (iii) using the Monotone Class Theorem. The last statement rewrites (iii) in terms of random variables. ■

Proof. [Of Theorem 34] For a fixed $N > 0$, consider a sequence of $\mathbb{G}^n$ predictable elementary processes of the form

$$H_t^m(\omega) = \sum_{i=1}^{k_m} h_i^m(\omega) 1_{]t_i^m, t_{i+1}^m]}(t),$$

null outside the fixed interval $[0, N]$ and with h_i^m $\mathcal{G}_{t_i^m}^n$-measurable and bounded, and $0 \leq t_1^m \leq \ldots \leq t_{k_m+1}^m < \infty$. Suppose that H^m converges to zero uniformly in (ω, t). By the Bichteler-Dellacherie's Theorem, we need to prove that $\lim_{m \to \infty} J_X(H^m) = 0$ in L^0.

For each fixed $p \geq 0$, define

$$J_t^{m,p} = \sum_{i=1}^{k_m} h_i^m \mathbf{1}_{\{\tau_{p+1}^n > t_i^m\}} \mathbf{1}_{]t_i^m, t_{i+1}^m]}.$$

By Lemma 38 (iv) there are $\mathcal{G}_{t_i^m}^{n,p}$-measurable random variables $j_i^{m,p}$ such that

$$h_i^m \mathbf{1}_{\{\tau_{p+1}^n > t_i^m\}} = j_i^{m,p} \mathbf{1}_{\{\tau_{p+1}^n > t_i^m\}} \text{ a.s.}$$

The latter quantity is $\mathcal{G}_{t_i^m}^{n,p+1}$-measurable, implying that the processes $J^{m,p}$ are $\mathbb{G}^{n,p+1}$ predictable. Moreover, since the h_i^m tend to zero uniformly in i and ω, it follows that for each fixed p, the processes $J^{m,p}$ tend to zero uniformly in (ω, t) except on a nullset.

Now, it is clear that $J^{m,p} = H^m$ on $\{\tau^n_{p+1} > N\}$. For any fixed $\varepsilon > 0$ we thus have:

$$\begin{aligned} P(|J_X(H^m)| \geq \varepsilon) &\leq E(\mathbf{1}_{\{|J_X(H^m)|\geq\varepsilon\}}\mathbf{1}_{\{\tau^n_{p+1}>N\}}) + P(\tau^n_{p+1} \leq N) \\ &= E(\mathbf{1}_{\{|J_X(J^{m,p})|\geq\varepsilon\}}\mathbf{1}_{\{\tau^n_{p+1}>N\}}) + P(\tau_{p+1} \leq N). \end{aligned}$$

We begin with the first term. Since M is a $\mathbb{G}^{n,p+1}$ semimartingale by Condition 4 (i) and since $J^{m,p}$ is $\mathbb{G}^{n,p+1}$ predictable and a.s. uniformly convergent to zero, it follows that $J_X(J^{m,p})$ tends to zero in probability by the Dominated Convergence Theorem for stochastic integrals. Therefore, as $m \to \infty$,

$$E(\mathbf{1}_{\{|J_X(J^{m,p})|\geq\varepsilon\}}\mathbf{1}_{\{\tau^n_{p+1}>N\}}) \to 0.$$

Hence, for each p,

$$\lim_{m\to\infty} P(|J_X(H^m)| \geq \varepsilon) \leq P(\tau^n_{p+1} \leq N).$$

Let $p \to \infty$ and use that $\lim_{p\to\infty} \tau^n_p = \infty$ a.s. to see that the left side above is indeed zero. This proves the result. ■

We focus now on Condition 4 (i). We use Theorem 27 to guarantee that this condition is satisfied under suitable assumptions. Introduce the intermediate filtrations $\mathbb{H}^{n,k} = (\mathcal{H}^{n,k}_t)_{0\leq t\leq T}$ where

$$\mathcal{H}^{n,k}_t = \bigcap_{u>t} \mathcal{F}_u \vee \sigma(\tau^n_i, \varepsilon_{n,i}, 1 \leq i \leq k)$$

As in Theorem 27, we assume that the following holds.

Assumption 5 *For each $n \geq 1$ and each $k \geq 1$, there exists an $\mathbb{H}^{n,k}$ predictable finite variation process $B^{n,k}$ such that $M - B^{n,k}$ is a $\mathbb{H}^{n,k}$ local martingale.*

In particular, Assumption 5 holds when the following is satisfied.

Assumption 6 (Initial times Assumption) *For each $n \geq 1$ and $k \geq 1$, $(\tau^n_i, 1 \leq i \leq k)$ satisfies Jacod's criterion.*

In this case one can enlarge initially with the random vector $(\tau^n_i, 1 \leq i \leq k)$ to obtain $\tilde{\mathcal{H}}_t = \bigcap_{u>t} \mathcal{F}_t \vee \sigma(\tau^n_i, 1 \leq i \leq k)$, and then expand initially $\tilde{\mathbb{H}}$ with the random vector $(\varepsilon^n_i, 1 \leq i \leq k)$ which always satisfies Jacod's countable expansion criterion since $(\varepsilon^n_i, 1 \leq i \leq k)$ takes values in the finite set $\{-1, 1\}^k$. The processes $B^{n,i}$ are then given by

$$B^{n,i}_t = \int_0^t \frac{d\langle M, p^{n,i}(\theta)\rangle_s}{p^{n,i}_{s-}(\theta)}\Bigg|_{\theta=(\varepsilon_{n,p},\tau_{n,p})_{1\leq p\leq i}}$$

where $p^{n,i}$ is the $\mathbb{F}$ conditional density of the random vector $(\varepsilon_{n,p}, \tau_{n,p})_{1\leq p\leq i}$. Let $Z^{n,i}_t = P(\tau^n_i > t \mid \mathcal{H}^{n,i}_t)$ and $\mu^{n,i}$ is the martingale part in its Doob-Meyer decomposition and $J^{n,i}$ the dual predictable projection of $\Delta M_{\tau^n_i} 1_{[\tau^n_i,\infty[}$ onto $\mathbb{H}^{n,i}$. Theorem 27 ensures that the following holds.

Theorem 35 *If Assumption 5 is satisfied then M is a $\mathbb{G}^{n,k}$ semimartingale with local martingale part $M_t - A_t^{n,k}$ where*

$$A_t^{n,k} = \sum_{i=0}^{k-1} \Big(\int_{t\wedge\tau_i^n}^{t\wedge\tau_{i+1}^n} dB_s^{n,i} + \int_{t\wedge\tau_i^n}^{t\wedge\tau_{i+1}^n} \frac{d\langle M, \mu^{n,i+1}\rangle_s + dJ_s^{n,i+1}}{Z_{s^-}^{n,i+1}} \Big) + \int_{t\wedge\tau_k^n}^{t} dB_s^{n,k} \qquad (1.11)$$

Condition 4 (i) is therefore satisfied. Moreover if either Assumption 4 or Condition 4 (ii) holds, then M is a $\mathbb{G}^n$ semimartingale for each n.

Proof. The first claim follows directly from Theorem 27 in the special case where the random times are ranked. Now, if Assumption 4 is satisfied, it follows from Theorem 34 that for each $n \geq 1$, M is a $\mathbb{G}^n$ semimartingale. If Condition 4 (ii) holds, we combine Lemma 37 and Theorem 22 to get the same conclusion. ■

Before stating the main result of this section, we provide sufficient conditions on X that guarantee that Assumptions 4 and 5 (and even Assumption 6) are satisfied.

Lemma 39 *Assume that X is a continuous non explosive increasing process starting at zero and such that $\lim_{t\to\infty} X_t = \infty$. Assume also that the conditional probabilities*

$$P_t^n(u_1, \dots, u_p) = P(X_{u_1} \leq \varepsilon_n, \dots, X_{u_p} \leq p\varepsilon_n \mid \mathcal{F}_t)$$

are differentiable w.r.t $(u_1, \dots, u_p)$, for each $n, p \geq 1$ and $0 \leq t \leq T$. Then Assumptions 4 and 6 (and hence 5) are satisfied.

Proof. First, since X is continuous and increasing, $X_{\tau_p^n} = p\varepsilon_n$ for all $n, p \geq 1$, therefore

$$\tau_p^n = \inf\{t > \tau_{p-1}^n,\ X_t \geq p\varepsilon_n\} = \inf\{t > 0,\ X_t \geq p\varepsilon_n\}$$

where the second equality follows from the fact that $X_s \leq (p-1)\varepsilon_n$, on $\{s \leq \tau_{p-1}^n\}$. Since X does not explode in finite time and $\lim_{t\to\infty} X_t = \infty$, it follows that $\lim_{p\to\infty} \tau_p^n = \infty$. This proves that Assumption 4 is satisfied.

Now notice that $\{\tau_p^n > u\} = \{\sup_{0\leq s\leq u} X_s < p\varepsilon_n\}$. Therefore, for each $n, p \geq 1$ and $u_1 \leq \dots \leq u_p$,

$$P(\tau_1^n > u_1, \dots, \tau_p^n > u_p \mid \mathcal{F}_t) = P(\sup_{0\leq s\leq u_1} X_s \leq \varepsilon_n, \dots, \sup_{0\leq s\leq u_p} X_s \leq p\varepsilon_n \mid \mathcal{F}_t)$$

But since X is increasing, $\sup_{0\leq s\leq u_k} X_s = X_{u_k}$, therefore

$$P(\tau_1^n > u_1, \dots, \tau_p^n > u_p \mid \mathcal{F}_t) = P(X_{u_1} \leq \varepsilon_n, \dots, X_{u_p} \leq p\varepsilon_n \mid \mathcal{F}_t) = P_t^n(u_1, \dots, u_p)$$

is differentiable w.r.t $(u_1, \dots, u_p)$. This proves that $(\tau_1^n, \dots, \tau_p^n)$ has an $\mathbb{F}$ conditional density and Assumption 6 is satisfied for each $p \geq 1$ (hence Assumption 5 too is satisfied). Finally note that since X is increasing, $\varepsilon_{n,p} = \varepsilon_n$, for all $n, p \geq 1$ are deterministic constants and the filtration $\mathbb{H}$ and $\tilde{\mathbb{H}}$ are the same. ■

1.5.1.2 Condition 3 (ii)

In this subsection, Assumptions 4 and 5 remain in force. We need to obtain the explicit canonical decomposition of M when viewed in $\mathbb{G}^n$ to be able to check Condition 3 (ii). Note that thanks to Assumption 4, $\lim_{k\to\infty} A_t^{n,k}$ is well defined, for each $n \geq 1$ and $0 \leq t \leq T$. Let A^n be this a.s. limit

$$A_t^n = \sum_{i=0}^{\infty} \int_{t\wedge\tau_i^n}^{t\wedge\tau_{i+1}^n} dB_s^{n,i} + \int_{t\wedge\tau_i^n}^{t\wedge\tau_{i+1}^n} \frac{d\langle M, \mu^{n,i+1}\rangle_s + dJ_s^{n,i+1}}{Z_{s^-}^{n,i+1}} \tag{1.12}$$

Of course, for each n and t, the sum above contains a.s. a finite number of terms. For A^n to be the $\mathbb{G}^n$ predictable finite variation term in the canonical decomposition of M, some conditions are needed.

Assumption 7 *The processes X and M are such that*

(i) The processes $A^{n,k}$ and A^n have integrable total variations.

(ii) $\lim_{p\to\infty} E\big(1_{\{\tau_p^n\leq t\}}|A_t^{n,p} - A_t^n|\big) = 0$*, for each $n \geq 1$ and t.*

where $A^{n,k}$ are defined in (1.11) and A^n are defined in (1.12).

Note that Condition 4 (ii) implies Assumption 7 (i) by Theorem 22 (ii). The following theorem is the main result of this section.

Theorem 36 *Assume that X and M are such that Assumptions 4, 5 and 7 are satisfied. Then*

(i) M is a $\mathbb{G}^n$ semimartingale with martingale part $M - A^n$, where A^n is given in (1.12).

(ii) If moreover Condition 3(ii) is satisfied, i.e. if $E(\int_0^T |dA_s^n|) \leq K$ for some K and all n, then M is a $\mathbb{G}$ special semimartingale.

Proof. Since Assumptions 4 and 5 are satisfied, it follows from Theorem 35 that M is a $\mathbb{G}^n$ semimartingale for each n. Under Assumption 7, it is clearly a $\mathbb{G}^n$ special semimartingale and we can obtain explicitly its canonical decomposition. Let $s \leq t$ and let U_s^n be a $\mathcal{G}_s^n$ measurable and bounded random variable. Since $\lim_{p\to\infty} \tau_p^n = \infty$ for each n, it follows from the monotone convergence theorem that

$$E((M_t - A_t^n)U_s^n) = \lim_{p\to\infty} E((M_t - A_t^n)U_s^n 1_{\{\tau_p^n > s\}})$$

From Lemma 38 (iv), for each $p \geq 1$, there exists a $\mathcal{G}_s^{n,p-1}$ measurable and bounded random variable $U_s^{n,p-1}$ such that

$$U_s^n 1_{\{\tau_p^n > s\}} = U_s^{n,p-1} 1_{\{\tau_p^n > s\}}$$

Therefore,

$$\begin{aligned}
E((M_t - A_t^n)U_s^n) &= \lim_{p\to\infty} E((M_t - A_t^n)U_s^{n,p-1}1_{\{\tau_p^n>s\}}) \\
&= \lim_{p\to\infty} E(U_s^{n,p-1}1_{\{\tau_p^n>s\}}E(M_t - A_t^n \mid \mathcal{G}_s^{n,p})) \\
&= \lim_{p\to\infty} E\Big(U_s^{n,p-1}1_{\{\tau_p^n>s\}}\big(E(M_t - A_t^{n,p} \mid \mathcal{G}_s^{n,p}) - E(A_t^n - A_t^{n,p} \mid \mathcal{G}_s^{n,p})\big)\Big)
\end{aligned}$$

Since $M - A^{n,p}$ is a $\mathbb{G}^{n,p}$ local martingale and $\sup_{0\le s\le T}|M_s|$ is integrable and $A^{n,p}$ has integrable total variation, it follows that $M - A^{n,p}$ is a $\mathbb{G}^{n,p}$ martingale. Therefore

$$\begin{aligned}
E((M_t - A_t^n)U_s^n) &= \lim_{p\to\infty} E\Big(U_s^{n,p-1}1_{\{\tau_p^n>s\}}\big(M_s - A_s^{n,p}) - E(A_t^n - A_t^{n,p} \mid \mathcal{G}_s^{n,p})\big)\Big) \\
&= \lim_{p\to\infty} E\Big(U_s^{n,p-1}1_{\{\tau_p^n>s\}}\big(M_s - A_s^{n,p} - A_t^n + A_t^{n,p}\big)\Big) \\
&= \lim_{p\to\infty} E\big(U_s^n 1_{\{\tau_p^n>s\}}(M_s - A_s^n)\big) \\
&+ \lim_{p\to\infty} \Big(E\big(U_s^n 1_{\{\tau_p^n>s\}}(A_s^n - A_s^{n,p})\big) + E\big(U_s^n 1_{\{\tau_p^n>s\}}(A_t^{n,p} - A_t^n)\big)\Big)
\end{aligned}$$

It follows from the definition of $A^{n,p}$ and A^n that $A_t^{n,p} = A_t^n$ on the set $\{\tau_p^n > t\}$. Therefore

$$E((M_t - A_t^n)U_s^n) = E((M_s - A_s^n)U_s^n) + \lim_{p\to\infty} E\big(U_s^n 1_{\{s<\tau_p^n\le t\}}(A_t^{n,p} - A_t^n)\big)$$

where we also used a monotone convergence theorem and the fact that $\lim_{p\to\infty}\tau_p^n = \infty$. By Assumption 7, this converges to $E((M_s - A_s^n)U_s^n)$, which proves that $M - A^n$ is a $\mathbb{G}^n$ martingale. For (ii), just use that $\mathcal{G}_t^n \xrightarrow{w} \mathcal{G}_t$ together with Theorem 22 to conclude that M is a $\mathbb{G}$ semimartingale. ■

However, it turned out to be difficult to find practical examples where the times τ_p^n are initial. In the next subsection, we do the same analysis under the assumption that these hitting times are honest.

1.5.2 Filtration expansion results based on a honest times assumption

We assume that the continuous process X is increasing to infinity and we define as in Lemma 39, the sequences of random times $(\tau_p^n)_{p\ge 0}$ to be

$$\tau_p^n = \inf\{t \ge 0, X_t \ge p\varepsilon_n\}$$

Since X is increasing to infinity, $(\tau_p^n)_{p\ge 0}$ satisfies Assumption 4 for each $n \ge 1$. Define the sequence of processes

$$X_t^n = \sum_{p=0}^{\infty} 1_{\{\tau_p^n\le t<\tau_{p+1}^n\}} X_{\tau_p^n} = \varepsilon_n \sum_{p=0}^{\infty} p 1_{\{\tau_p^n\le t<\tau_{p+1}^n\}}$$

and $\mathbb{G}^n$ (resp. $\mathbb{G}$) the progressive expansion of $\mathbb{F}$ with X^n (resp. X). Let $\mathbb{G}^{\tau^n}$ be the smallest filtration containing $\mathbb{F}$ and that makes all $(\tau_p^n)_{p\geq 1}$ stopping times. Clearly $\mathbb{G}^n = \mathbb{G}^{\tau^n}$. We make now the following assumption.

Assumption 8 (Honest times assumption) *For each $n \geq 1$, the sequence $(\tau_p^n)_{p\geq 0}$ is an increasing sequence of $\mathbb{F}$ honest times such that $\tau_0^n = 0$ and $\sup_p \tau_p^n = \infty$.*

Putting Theorem 10 and Theorem 22 together, we obtain the following result, which is the main result of this section.

Theorem 37 *Suppose Assumption 8 holds. Let M be an $\mathbb{F}$ martingale such that $\sup_{0\leq s\leq T}|M_s|$ is integrable. If $E(\int_0^T |dA_s^n|) \leq K$ for some K and all $n \geq 1$, then M is a $\mathbb{G}$ semimartingale. Here A^n is defined in equation (1.13).*

Proof. By Theorem 10, under assumption 8, M is a $\mathbb{G}^{\tau^n}$ semimartingale and $M - A^n$ is a $\mathbb{G}^{\tau^n}$ local martingale, where

$$A_t^n = \sum_{p=0}^{\infty} \int_0^t 1_{\{\tau_p^n < s \leq \tau_{p+1}^n\}} \frac{1}{Z_{s^-}^{n,p+1} - Z_{s^-}^{n,p}} d\langle M, M^{n,p+1} - M^{n,p}\rangle_s \tag{1.13}$$

where $Z^{n,p}$ is the $\mathbb{F}$ optional projection of τ_p^n and $M^{n,p}$ is the martingale part in its Doob-Meyer decomposition. Therefore, M a $\mathbb{G}^n$ semimartingale. Now, $\mathcal{G}_t^n$ converges weakly to $\mathcal{G}_t$, for each t and $E(\int_0^T |dA_s^n|) \leq K$ for some K and all $n \geq 1$ by assumption. Theorem 22 allows to conclude. ■

1.5.3 Insider models with arbitrage: Jeulin's example and extensions

There is generally no link between the absolute continuity of A^n and the absolute continuity of A. We first suggest a toy convincing example, before constructing more elaborate examples that satisfy the assumptions of the previous section. This is important from a practical point of view, since a singular finite variation term would introduce *extreme* arbitrage opportunities : there will be no equivalent local martingale measure under which the price process is a local martingale. We check first that Jeulin's example (expansion of the natural filtration of a Bessel 3 process with its remaining infimum, see [89]) falls within the class of examples that Theorem 37 allows to deal with and generalize this result to the expansion of the natural filtration of a transient diffusion with its remaining infimum. In these examples, the finition variation part in the extended filtration of a given Brownian motion of the base filtration becomes singular w.r.t Lebesgue measure, which can be interpreted in financial terms as riskless insider trading.

Let us start with the following easy example where the absolute continuity of A^n for each $n \geq 1$ is not inherited by A. Let N be a Poisson process with intensity λ and $\mathbb{F}$ its natural filtration. Let T be the first jump time of N.

Lemma 40 *T does not satisfy Jacod's criterion and its $\mathcal{F}_s$ conditional density is given by*

$$p_s(dx) = 1_{\{x \le s\}}\delta_T(dx) + 1_{\{x>s\}}\lambda e^{-\lambda(x-s)}$$

Proof. Since T is an $\mathbb{F}$ stopping time, $P(T \le t \mid \mathcal{F}_s) = 1_{\{T\le t\}}$ if $t \le s$. Let $t \ge s$.

$$\begin{aligned} P(T \le t \mid \mathcal{F}_s) &= P(N_t \ge 1 \mid \mathcal{F}_s) = 1_{\{N_s \ge 1\}} + 1_{\{N_s=0\}}P(N_t \ge 1 \mid \mathcal{F}_s) \\ &= 1_{\{T \le s\}} + 1_{\{T>s\}}P(N_{t-s} \ge 1) = 1_{\{T\le s\}} + 1_{\{T>s\}}(1 - e^{-\lambda(t-s)}) \end{aligned}$$

Therefore $p_s(dx) = 1_{\{x\le s\}}\delta_T(dx) + 1_{\{x>s\}}\lambda e^{-\lambda(x-s)}$. ∎

Let $T_n = T + U_n$ with $U_n \to 0$ a.s. and such that U_n has a density g_n. Then T_n satisfies Jacod's criterion with $p^n_s(t)dt = E(g_n(t-T) \mid \mathcal{F}_s)dt$. From Lemma 40,

$$p^n_s(t) = 1_{\{T\le s\}}g_n(t-T) + 1_{\{T>s\}}\int_s^\infty \lambda e^{-\lambda(u-s)}g_n(t-u)du$$

Since $\tilde{N}_t = N_t - \lambda t$ is an $\mathbb{F}$ martingale, it follows that

$$\tilde{N}_t - \int_0^t \frac{\langle N, p^n(u)\rangle_s}{p^n_{s-}(u)}\Big|_{u=T_n}$$

is an H^n martingale. By Kunita Watanabe's inequality, dA^n is absolutely continuous w.r.t Lebesgue measure, where $A^n_t = \int_0^t \frac{\langle N,p^n(u)\rangle_s}{p^n_{s-}(u)}\Big|_{u=T_n}$. However N becomes an $\mathbb{H}$ semimartingale with singular compensator A such that A has integrable total variation on compact time intervals.

Lemma 41 *$M_t := N_t - \lambda(t-T)1_{\{T\le t\}} - 1_{\{T\le t\}}$ is an $\mathbb{H}$ martingale.*

Proof. Let $s \le t$, f_s an $\mathcal{F}_s$ measurable bounded and h a Borel bounded function.

$$\begin{aligned} E((N_t - N_s)f_s h(T)) &= E((N_t - N_s)f_s h(T)1_{\{T\le t\}}) \\ &= E((N_t - N_s)f_s h(T)(1_{\{T\le s\}} + 1_{\{s<T\le t\}})) \\ &= E(f_s h(T)\lambda(t-s)1_{\{T\le s\}}) + E(f_s h(T)1_{\{s<T\le t\}}(1 + E(N_t - N_T \mid \mathcal{F}_T))) \\ &= E\Big(f_s h(T)\big(\lambda(t-s)1_{\{T\le s\}} + 1_{\{s<T\le t\}}(1+\lambda(t-T))\big)\Big) \end{aligned}$$

A Monotone Class argument yields

$$E(N_t \mid \mathcal{H}_s) = N_s + \lambda(t-s)1_{\{T\le s\}} + (1+\lambda(t-T))1_{\{s<T\le t\}}$$

Writing $1_{\{s<T\le t\}} = 1_{\{T\le t\}} - 1_{\{T\le s\}}$ finally yields $E(M_t \mid \mathcal{H}_s) = M_s$. ∎

Therefore $A_t = \lambda(t-T)1_{\{T\le t\}} + 1_{\{T\le t\}}$ and dA is singular. We turn now to elaborate examples satisfying Assumption 8.

1.5.3.1 Expansion of the natural filtration of a Bessel 3 process with its remaining infimum

Let Z be a Bessel 3 process, $\mathbb{F}$ its natural filtration, $X_t = \inf_{s>t} Z_s$ and $\mathbb{G}$ the progressive expansion of $\mathbb{F}$ with X. This example has been studied in detail by both Jeulin and Pitman using different techniques. Let $B_t = Z_t - \int_0^t \frac{ds}{Z_s}$. It is a classical result that B is an $\mathbb{F}$ Brownian motion. Using Williams' path decomposition for Brownian motion, Pitman proves that $B_t - (2X_t - \int_0^t \frac{ds}{Z_s})$ is a $\mathbb{G}$ Brownian motion. Using filtration expansion results, Jeulin proves in [89] the $\mathbb{G}$ semimartingale property of B and provides its decomposition simultaneously. However, his technique is hard to generalize. Our approach allows to prove the semimartingale property of B, without the explicit decomposition, however, it has the merit that it can be used in much more general settings as described in Theorem 37.

Theorem 38 *The process B is a $\mathbb{G}$ special semimartingale.*

Proof. First X is continuous increasing to infinity since $\lim_{t\to\infty} Z_t = \infty$ (see Lemma 6.20 in Jeulin). Therefore $\tau_p^n = \inf\{t, X_t \geq p\varepsilon_n\}$ is increasing to infinity, for each n. Now, with $Y_t = 2X_t - Z_t$, we have $X_t = \sup_{s\leq t} Y_s$, so that

$$\tau_p^n = \inf\{t, X_t \geq p\varepsilon_n\} = \inf\{t, Y_t \geq p\varepsilon_n\} = \sup\{t, Z_t = p\varepsilon_n\}$$

Therefore $(\tau_p^n)_{p\geq 1}$ satisfies Assumption 8. We compute now A^n as of equation 1.13. It is a classical result that

$$Z_t^{n,p} = P(\tau_p^n > t \mid \mathcal{F}_t) = 1 \wedge \frac{p\varepsilon_n}{Z_t}$$

and Tanaka's formula implies that

$$M_t^{n,p} = 1 - p\varepsilon_n \int_0^t 1_{\{Z_s > p\varepsilon_n\}} \frac{dB_s}{Z_s^2}$$

Therefore, and since on $\{\tau_p^n < s\}$, $Z_s > p\varepsilon_n$,

$$\begin{aligned}
A_t^n &= \sum_{p=0}^{\infty} \int_0^t 1_{\{\tau_p^n < s \leq \tau_{p+1}^n\}} \frac{d\langle B, M^{n,p+1} - M^{n,p}\rangle_s}{Z_s^{n,p+1} - Z_s^{n,p}} \\
&= \sum_{p=0}^{\infty} \int_0^t 1_{\{\tau_p^n < s \leq \tau_{p+1}^n\}} \frac{\frac{p\varepsilon_n}{Z_s^2} - 1_{\{Z_s > (p+1)\varepsilon_n\}} \frac{(p+1)\varepsilon_n}{Z_s^2}}{\frac{(p+1)\varepsilon_n}{Z_s} 1_{\{(p+1)\varepsilon_n < Z_s\}} + 1_{\{(p+1)\varepsilon_n \geq Z_s\}} - \frac{p\varepsilon_n}{Z_s}} ds \\
&= \sum_{p=0}^{\infty} \int_0^t 1_{\{\tau_p^n < s \leq \tau_{p+1}^n\}} \frac{p\varepsilon_n - 1_{\{Z_s > (p+1)\}}(p+1)\varepsilon_n}{Z_s\big((p+1)\varepsilon_n 1_{\{(p+1)\varepsilon_n < Z_s\}} - p\varepsilon_n\big) + Z_s^2 1_{\{(p+1)\varepsilon_n \geq Z_s\}}} ds
\end{aligned}$$

Therefore

$$\begin{aligned}
A_t^n &= \sum_{p=0}^{\infty} \int_0^t 1_{\{\tau_p^n < s \leq \tau_{p+1}^n\}} \Big(1_{\{(p+1)\varepsilon_n \geq Z_s\}} \frac{p\varepsilon_n}{-Z_s p\varepsilon_n + Z_s^2} + 1_{\{(p+1)\varepsilon_n < Z_s\}} \frac{-1}{Z_s}\Big) ds \\
&= \sum_{p=0}^{\infty} \int_0^t 1_{\{\tau_p^n < s \leq \tau_{p+1}^n\}} \Big(-\frac{1}{Z_s} + 1_{\{(p+1)\varepsilon_n \geq Z_s\}} \frac{1}{Z_s - p\varepsilon_n}\Big) ds
\end{aligned}$$

Fubini's theorem implies finally that

$$A_t^n = -\int_0^t \frac{ds}{Z_s} + \sum_{p=0}^{\infty} \int_0^t 1_{\{\tau_p^n < s\}} 1_{\{(p+1)\varepsilon_n \geq Z_s\}} \frac{1}{Z_s - p\varepsilon_n} ds$$

where we also used $1_{\{s \leq \tau_{p+1}^n\}} 1_{\{Z_s \leq (p+1)\varepsilon_n\}} = 1_{\{Z_s \leq (p+1)\varepsilon_n\}}$. Now

$$E(\sum_{p=0}^{\infty} \int_0^t 1_{\{\tau_p^n < s\}} 1_{\{(p+1)\varepsilon_n \geq Z_s\}} \frac{1}{Z_s - p\varepsilon_n} ds) = \sum_{p=0}^{\infty} E(\int_0^t 1_{\{\tau_p^n < s\}} 1_{\{(p+1)\varepsilon_n \geq Z_s\}} \frac{1}{Z_s - p\varepsilon_n}) ds$$

$$= \sum_{p=0}^{\infty} E(\int_0^t \frac{1}{Z_s} 1_{\{Z_s > p\varepsilon_n\}} 1_{\{Z_s \leq (p+1)\varepsilon_n\}}) ds = E(\int_0^t \frac{1}{Z_s} \sum_{p=0}^{\infty} 1_{\{p\varepsilon_n < Z_s \leq (p+1)\varepsilon_n\}} ds) = E(\int_0^t \frac{ds}{Z_s})$$

where the second equality follows because the $\mathbb{F}$ optional projection of $1_{\{\tau_p^n \leq \cdot\}}$ is $(1 - \frac{p\varepsilon_n}{Z_t})^+$. It remains to use Theorem 37 to conclude. ∎

Using both our result and the explicit decomposition provided by Jeulin and Pitman, we see that although each A^n is absolutely continuous w.r.t Lebesgue measure, the finite variation term in the $\mathbb{G}$ decomposition of B is singular. In financial terms, this introduces arbitrage opportunities in the market where an insider discovers the extra information X_t progressively.

In this case of arbitrage opportunities, arising because Condition (i) of Theorem 33 is violated, Jarrow and Protter showed how one can explicitly construct an arbitrage opportunity. Following the techniques in [80], we show how this can be achieved. In order to do this, we recall the main result of their paper. They are able to explicitly construct arbitrage opportunities in the setting of the theorem below.

Theorem 39 (Jarrow and Protter) *Let $S_t = M_t + A_t$ be a continuous semimartingale. Then*

(i) There exists h predictable, and L continuous and adapted, such that

$$A_t = \int_0^t h_s d\langle M, M\rangle_s + L_t$$

where $t \to L_t$ and $t \to \langle M, M\rangle_t$ are singular measures on $\mathbb{R}^+$ a.s. Moreover the decomposition above is unique.

(ii) Assume furthermore that L is non-trivial and that there exists a unique Q equivalent to P such that $N_t = M_t + \int_0^t h_s d\langle M, M\rangle_s$ is a Q local martingale. Then there is an arbitrage opportunity.

In our case, S is the $\mathbb{F}$ Brownian motion B. So, of course, there are no arbitrage opportunities for S in $\mathbb{F}$. Now let us see what happens in the expanded filtration $\mathbb{G}$. The process $M_t = B_t - (2X_t - \int_0^t \frac{1}{Z_s} ds)$ is a $\mathbb{G}$ Brownian motion, and its quadratic variation is therefore $\langle M, M\rangle_t = t$ and $A_t = 2X_t - \int_0^t \frac{1}{Z_s} ds$. Therefore

$$h_t = -\frac{1}{Z_t}, \qquad N_t = M_t - \int_0^t \frac{1}{Z_s} ds \qquad \text{and} \qquad L_t = 2X_t$$

It is easy to see that there exists a unique equivalent martingale measure for N and the construction in [80] holds. That is, choose the contingent claim $H = L_T \in \mathcal{G}_T$. Following the lines of arguments of Jarrow and Protter, we show that the strategy $a_s = 1_{\{s \in supp(dL)\}}$ is an arbitrage opportunity for S is $\mathbb{G}$. In fact, it follows from Theorem 3.2 in [80] that there exists a self-financing strategy (j, b) (see [80] for a definition) such that

$$H = L_T = E^Q(L_T) + \int_0^T j_s dB_s$$

However we also have $L_T = 0 + \int_0^T a_s dL_s$. Moreover we have $\int_0^t a_s dN_s = 0$ for each $0 \leq t \leq T$, by construction of the process a. Therefore

$$H = L_T = 0 + \int_0^T a_s dB_s$$

which proves that the strategy a is an arbitrage opportunity.

1.5.3.2 The case of transient diffusions

Let R_t be a transient diffusion with values in $\mathbb{R}^+$, which has $\{0\}$ as entrance boundary. Let s be a scale function for R, which we can choose such that

$$\lim_{x \to 0} s(x) = -\infty \qquad \text{and} \qquad \lim_{x \to \infty} s(x) = 0$$

Let $\mathbb{F}$ be the natural filtration of R. Nikeghbali studied in [115] progressive filtration expansions of $\mathbb{F}$ with last exit times of such diffusions. Define X to be the remaining infimum of R, i.e. $X_t = \inf_{s>t} R_s$ and $\mathbb{G}$ the progressive expansion of $\mathbb{F}$ with X. We provide a sufficient condition for some $\mathbb{F}$ martingales to remain $\mathbb{G}$ semimartingales. The process $M_t = -s(R_t)$, which is well known to be an $\mathbb{F}$ local martingale will be of crucial interest. The next theorem is the main result of this section.

Theorem 40 *Let N be an $\mathbb{F}$ martingale such that $\sup_{0 \leq s \leq T} |N_s|$ and $\int_0^T \frac{|d\langle N, M\rangle_s|}{M_s}$ are integrable. Then N is a $\mathbb{G}$ semimartingale.*

Proof. Define the $\mathbb{F}$ honest random times $\sigma_y = \sup\{t \mid R_t = y\}$. Let $Y_t = 2X_t - R_t$, then $X_t = \sup_{s \leq t} Y_s$ and the random times

$$\tau_p^n := \inf\{t \mid X_t \geq p\varepsilon_n\} = \inf\{t \mid Y_t \geq p\varepsilon_n\} = \sup\{t \mid R_t = p\varepsilon_n\} = \sigma_{p\varepsilon_n}$$

are $\mathbb{F}$ honest and $(\tau_p^n)_{p \geq 1}$ satisfies Assumption 8. To use Theorem 37, we need to compute A^n as defined in (1.13). It is proved in [115] that

$$P(\sigma_y > t \mid \mathcal{F}_t) = \frac{s(R_t)}{s(y)} \wedge 1 = 1 - \frac{1}{s(y)} \int_0^t 1_{\{R_u > y\}} dM_u + \frac{1}{2s(y)} L_t^{s(y)}$$

where $L^{s(y)}$ is the local time of $s(R)$ at $s(y)$. Introduce $Z_t^{n,p} = P(\tau_p^n > t \mid \mathcal{F}_t) = \frac{s(R_t)}{s(p\varepsilon_n)} \wedge 1$ and the martingale part in its Doob Meyer decomposition $M_t^{n,p} = 1 - \frac{1}{s(p\varepsilon_n)} \int_0^t 1_{\{R_u > p\varepsilon_n\}} dM_u$. Therefore

$$A_t^n = \sum_{p=0}^{\infty} \int_0^t 1_{\{\tau_p^n < s \le \tau_{p+1}^n\}} \frac{\frac{-1}{s((p+1)\varepsilon_n)} 1_{\{R_s > (p+1)\varepsilon_n\}} + \frac{1}{s(p\varepsilon_n)}}{1_{\{R_s \le (p+1)\varepsilon_n\}} + 1_{\{R_s > (p+1)\varepsilon_n\}} \frac{s(R_s)}{s((p+1)\varepsilon_n)} - \frac{s(R_s)}{s(p\varepsilon_n)}} d\langle N, M\rangle_s$$

where we used that $R_s > p\varepsilon_n$ on $\{\tau_p^n < s\}$, and the fact that $-s$ is positive non increasing by construction. Basic algebraic manipulations give

$$\begin{aligned} A_t^n &= \sum_{p=0}^{\infty} \int_0^t 1_{\{\tau_p^n < s \le \tau_{p+1}^n\}} \Big(\frac{1}{-s(R_s)} + 1_{\{R_s \le (p+1)\varepsilon_n\}} \big(\frac{1}{s(R_s)} + \frac{1}{s(p\varepsilon_n) - s(R_s)} \big) \Big) d\langle N, M\rangle_s \\ &= \int_0^t \frac{d\langle N, M\rangle_s}{M_s} + \sum_{p=0}^{\infty} \int_0^t 1_{\{\tau_p^n < s\}} 1_{\{R_s \le (p+1)\varepsilon_n\}} \frac{s(p\varepsilon_n)}{s(p\varepsilon_n) - s(R_s)} \frac{d\langle N, M\rangle_s}{s(R_s)} \end{aligned}$$

where the last equality follows from $1_{\{s \le \tau_{p+1}^n\}} 1_{\{R_s \le (p+1)\varepsilon_n\}} = 1_{\{R_s \le (p+1)\varepsilon_n\}}$. Now

$$\begin{aligned} E(\int_0^T |dA_s^n|) &\le E(\int_0^T |\frac{d\langle N, M\rangle_s}{M_s}|) \\ &\quad + \sum_{p=0}^{\infty} E(\int_0^T 1_{\{\tau_p^n < s\}} 1_{\{R_s \le (p+1)\varepsilon_n\}} |\frac{s(p\varepsilon_n)}{s(R_s)(s(p\varepsilon_n) - s(R_s))}| |d\langle N, M\rangle_s|) \\ &\le E(\int_0^T |\frac{d\langle N, M\rangle_s}{M_s}|) + \sum_{p=0}^{\infty} E(\int_0^T 1_{\{p\varepsilon_n < R_s \le (p+1)\varepsilon_n\}} \frac{1}{M_s} |d\langle N, M\rangle_s|) \\ &\le 2E\big(\int_0^T \frac{1}{M_s} d\langle N, M\rangle_s\big) \end{aligned}$$

where the second inequality uses that the $\mathbb{F}$ optional projection of $1_{\{\tau_p^n \le \cdot\}}$ is given by $\big(1 - \frac{s(R_s)}{s(p\varepsilon_n)}\big)^+$ and the monotonicity of s. Theorem 37 allows to conclude. ∎

ACKNOWLEDGMENT.

Section 1.3 is based on a joint work (see [95]) with Martin Larsson and Philip Protter. Sections 1.4 and 1.5 are based on a joint work (see [96] and [97]) with Philip Protter.

Chapter 2

Compensators of random times and credit contagion

Contents

2.1 Introduction

Recently, many researchers tried to reconcile the two main approaches in credit risk modeling : structural and reduced form models. Using structural models with incomplete information, many authors are able to obtain reduced form models in which the intensity of default is not given exogenously. This intensity is determined as a function of the firm's characteristics and the level of information that investors possess. Usually, in structural models, if the investor has complete information, he is able to predict the arrival of the default time. Reduced form models overcome this limitation specifying an exogenous default intensity which makes the default an unpredictable event. The advantages of modeling default times as totally inaccessible stopping times given the market's information set have long been recognized, see for instance [81], [45], [23], [56] and [57]. We summarize now briefly how this is achieved within reduced form models where default intensities are modeled directly. Given a market filtration $\mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$, the starting point is the description of a process $(\lambda_t)_{t\geq 0}$, the default intensity, such that

$$\mathbf{1}_{\{\tau\leq t\}} - \int_0^{t\wedge\tau} \lambda_s ds$$

is a $\mathbb{G}$ martingale, where the $\mathbb{G}$ stopping time τ represents the time of default. This raises the question of how to construct the stopping time τ when λ_t is already given. A widely used method is the *Cox construction*: start with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ and model $(\lambda_t)_{t\geq 0}$ as any nonnegative process adapted to $\mathbb{F}$. Then define

$$\tau = \inf\{t : \int_0^t \lambda_s ds > \Theta\},$$

where Θ is an exponentially distributed random variable independent of $\mathcal{F}_\infty$. Clearly τ is not an $\mathbb{F}$ stopping time; however, in the filtration $\mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$ obtained as the *progressive expansion* of $\mathbb{F}$ with τ,

$$\mathcal{G}_t = \bigcap_{u>t} \mathcal{F}_u \vee \sigma(\tau \wedge u),$$

it turns out that τ is a stopping time with intensity λ_t. Note also that some recent work ([88]) has been done to construct an explicit model of default time with a given survival probability. However, the main problem with reduced form models is that the arrival of default is not based on any characteristic of the firm's underlying credit quality. Using specifications of structural models where investors do not have complete information about the dynamics of the processes which trigger the firm's default (that is, relaxing the complete information assumption), many authors (see for instance [38], [23], [56], [57], [54],[55], and [59]) arrive to a framework which links both credit risk modeling approaches. In this chapter we take these studies a step further and try to fill the gap between reduced form and structural models from the perspective of credit contagion. For theoretical and practical studies of multi-name structural models, we refer the reader to [104] and [105] and the references therein. Information induced default contagion effect have been studied recently by researchers (see for instance [121], [103] and [27]) who tried to analyse and model credit contagion effects. The idea is quite simple: the default of one firm indirectly and instantaneously updates the market's knowledge of the state of the other firms, causing their default intensities to jump. Within reduced form models, the credit contagion effect is therefore taken into account by making the default intensity of a firm jump at the defaults of other firms. We look into different structural models under different sets of partial information and study the information induced credit contagion. Since many structural models can fall within the large class of models where the default times admit a conditional density (conditionally to the base filtration), we then focus our efforts on times satisfying this assumption and we are able to provide partial results reconciling structural models where the time admits a conditional density with reduced form models from the credit contagion effect's perspective.

The aim of the present chapter is therefore twofold. Our first goal is to study the information induced credit contagion effect within structural models, mainly through examples. After deriving the base model for a single firm, we use it as a building block for several multiple firms structural models and try to characterize the credit contagion effect within each one of them, under different levels of information. For simplicity, we usually study two firms models. We start with a non symmetric model, where the first firm value would depend only on some economy state factor Y, but the second firm depends not only on the economy but explicitly on the first firm value through its volatility. In the second model, the two firms will be co-dependent and the dependence arises through the drift terms. The base filtration $\mathbb{F}$ reveals the state factor Y (of course, it can be reduced to the trivial filtration if wanted), the firm values are not observable and investors observe only $\mathbb{F}$ and the times at which the two firms default. The default times are defined as the first hitting times of deterministic levels by the firms values. This type of model has also been studied in the work of Coculescu and al. [26], although not as intensively as in the present chapter. In order to prove existence and compute the default intensities, we use techniques from the filtering theory together with representation theorems in the progressively expanded filtrations. Here the intensities are expressed in terms of conditional expectations of quantities depending on the characteristics of the firm values, but are highly untractable. We then investigate two models with a different level of information, now the firms can be either unobservable or partially observable, but the default levels are essentially random. We introduce the contagion effect through the dependence of the threshold L^2 on the default time τ_1 of the first firm. The threshold L^2 is chosen of the form $a+b1_{\{\tau_1\leq T\}}$, where T is the fixed time horizon and a and b are two positive deterministic constants. This definition of L^2 is motivated by the fact that, financially speaking, the bankruptcy might occur because of the default event of firm 1 and although firm 2 is still in a relatively healthy economic situation. Note that this construction is inspired from [62] and [63], although only one firm is considered in that paper and no credit contagion issues are discussed. Even if this last model has the merit to be very tractable, we can usually compute in the general cases the intensities only in terms of conditional expectations. That these intensities jump at the default times of the other firms seems very likely but almost all examples are non tractable and use involved computations, and proving that these jumps occur in full generality is a hard task (apart for conditional independence models, where it is usually straightforward to prove that there is no credit contagion effect). Fortunately, many structural models would in practice benefit from the property that the default times admit an $\mathbb{F}$ conditional densities. Therefore, in the second section, we unify the approach under the standing assumption of existence of $\mathcal{F}_t$-conditional densities for the vector $\boldsymbol{\tau}$ of the random default times, i.e. $\mathbb{F}$ adapted processes $p_t(\boldsymbol{u})$, indexed by $\boldsymbol{u}\in\mathbb{R}_+^n$, such that

$$P(\boldsymbol{\tau}\in d\boldsymbol{u}\mid\mathcal{F}_t)=p_t(\boldsymbol{u})d\boldsymbol{u}.$$

In practice, this translates the difficulty from expressing directly the default intensities in terms of the parameters of the structural model to computing the conditional densities of the default times only. As seen in the previous chapter (see Assumption 1), this kind of assumption was first used by Jacod [72] in his paper on initial expansions of filtrations.

It has also been used later in the credit risk context by Jeanblanc and Le Cam [86] and more importantly, El Karoui, Jeanblanc and Jiao (see [40] and [41]) introduced this same assumption to study the case of ranked times. We use the techniques in [40] and [41] to extend the filtration expansion approach described briefly above to the case where there is an arbitrary number of default times $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_n)$ under the conditional density assumption. Moreover, we wish to carry out the analysis without imposing any restrictions on the ordering of the individual times. Concerning the passage from the ranked to the non-ranked case, the difficulty is mainly notational since the methods have already been developed by El Karoui, Jeanblanc and Jiao. The case with ranked defaults, that has been treated in [40], [41] and [53], is sufficient for certain applications, such as the pricing of defaultable bonds or other default sensitive securities. There are, however, situations where this is inadequate. We will focus on a first example from risk management, where the distribution of securities prices at some future time is of interest, for instance for value-at-risk or expected shortfall computations. The use of ranked default times can then yield highly misleading conclusions, as we will demonstrate. As in the case of one default time, we start with a base filtration $\mathbb{F}$ and model the market filtration $\mathbb{G}$ as the progressive expansion of $\mathbb{F}$ with the whole vector $\boldsymbol{\tau}$, i.e.

$$\mathcal{G}_t = \bigcap_{u>t} \mathcal{F}_u \vee \sigma(\tau_i \wedge u : i = 1, \ldots, n).$$

In the first parts of section 2.3, we slightly extend the results in [40] and [41], to the multiple non-ranked times case. Quantities such as $\mathbb{G}$ default intensities and pricing formulas will be given in terms of $p_t(\boldsymbol{u})$ and are needed for a subsequent study on risk management and credit contagion. It is desirable to model these densities directly, and we will discuss how this can be done. Analogously to the one-default case described previously, there is the issue of constructing the actual vector $\boldsymbol{\tau}$, given that a model for $p_t(\boldsymbol{u})$ has been chosen. We will show that this can always be done. Moreover, we provide a general scheme for simulating joint realizations of $p_t(\boldsymbol{u})$ and $\boldsymbol{\tau}$. We then point out again that the information induced default contagion effect (see for instance [121], [103] and [27]), where the default of one firm indirectly and instantaneously updates the market's knowledge of the state of the other firms, causing their default intensities to jump, naturally emerges in the framework of this section and we study this phenomenon for some particular form of conditional densities. Finally, we conclude this section by providing a toy example where the reconciliation between structural and reduced form models from the perspective of credit contagion is done explicitly.

2.2 Information induced credit contagion in structural models

In this section, we study the information induced credit contagion effect within structural models, mainly through examples. After deriving the base model for a single firm, we use it as a building block for several multiple firms structural models and try to characterize the credit contagion effect within each one of them, under different levels of information. As we will see, almost all examples are non tractable and computationally very involved. The aim here is to try to check *à la main* if the default intensity of a given firm (we will prove that they exist in the models we consider) jumps at the default times of the other firms. We first define what we explicitly mean by information induced credit contagion.

Definition 9 (Information induced credit contagion effect) *Given a filtration $\mathbb{F}$ and its progressive expansion $\mathbb{G}$ with random times $\tau_1, \ldots, \tau_n$, there is an information induced credit contagion effect for firm i at the default time of firm j if and only if the $\mathbb{G}$ default intensity λ^i jumps at the default time τ_j with positive probability. If this probability is one, we will say that there is an a.s. information induced credit contagion effect.*

To clarify ideas, take for now the particular case of two firms. This effect is described as *information induced* credit contagion as it arises due to the fact that the arrival of one default helps localize the other default time in the support of its conditional law. The interpretation is that the default of the first firm indirectly gives information about the state of the second firm, and vice versa. See [121] for a discussion of this phenomenon.

We first study a simplified model for a single firm, that would be generalized in the next sections.

2.2.1 The base model for a single time

We consider the following simplified model for a single firm:

$$\begin{aligned} dX_t &= X_t \sigma(Y_t) dB_t \\ dY_t &= Y_t dW_t \end{aligned}$$

with $X_0 = Y_0 = 1$, $\tau = \inf\{t : X_t \le \alpha\}$, for some fixed $\alpha \in (0,1)$ and σ a continuous funtion. Let $\mathbb{H}$ be the natural filtration of the two-dimensional Brownian motion (B, W), $\mathbb{F}$ the natural filtration of W, and let $\mathbb{G}$ be the progressive enlargement of $\mathbb{F}$ with τ.

2.2.1.1 The $\mathbb{F}$ conditional survival probability

Our goal in this subsection is to compute $P(\tau > t \mid \mathcal{F}_t)$. First define

$$F(t) = P\left(\inf_{s \le t} \beta_s - \frac{1}{2}s \le \ln \alpha\right), \tag{2.1}$$

where β is a Brownian motion. We then have the following result.

Lemma 42 *The conditional survival probability of τ given $\mathcal{F}_t$ is given by*

$$Z_t = P(\tau > t \mid \mathcal{F}_t) = 1 - F\left(\int_0^t \sigma(Y_s)^2 ds\right).$$

Proof. Define the functional $h : L[0,t] \times \Omega \to \mathbb{R}$ by

$$h(f) = h(f,\omega) = \mathbf{1}_{\{\inf_{s\leq t}(f\cdot B)_s(\omega) - \frac{1}{2}\int_0^s f(u)^2 du > \ln \alpha\}}.$$

For each fixed t, path continuity gives

$$\{\tau > t\} = \{\inf_{s\leq t} X_s > \alpha\} = \left\{\inf_{s\leq t}\int_0^s \sigma(Y_u)dW_u - \frac{1}{2}\int_0^s \sigma(Y_u)^2 du > \ln\alpha\right\}.$$

Since Y is independent of B, we get

$$P(\tau > t \mid \mathcal{F}_t) = E(h(\sigma(Y)) \mid \mathcal{F}_t) = E[h(f)]_{|f=\sigma(Y)},$$

To compute $E[H(f)]$, note that by the Dubins-Schwarz theorem there is a Brownian motion β, possibly on an extended probability space, such that $\int_0^t f(u)dW_u = \beta_{\int_0^t f(u)^2 du}$. With F defined in (2.1) we thus have

$$\begin{aligned}
E[h(f)] &= P\left(\inf_{s\leq t}\int_0^s f(u)dW_u - \frac{1}{2}\int_0^s f(u)^2 du > \ln\alpha\right) \\
&= P\left(\inf_{s\leq t}\beta_{\int_0^s f(u)^2 du} - \frac{1}{2}\int_0^s f(u)^2 du > \ln\alpha\right) \\
&= P\left(\inf_{s\leq \int_0^t f(u)^2 du}\beta_s - \frac{1}{2}s > \ln\alpha\right) \\
&= 1 - F\left(\int_0^t f(u)^2 du\right),
\end{aligned}$$

where the last equality used that f is deterministic. This finishes the proof. ∎

2.2.1.2 The $\mathbb{G}$ compensator

The function F is increasing, as is the process $(\int_0^t \sigma(Y_s)^2 ds)_{t\geq 0}$. Hence $(Z_t)_{t\geq 0}$ is decreasing. Continuity and absolute continuity with respect to Lebesgue measure of Z_t holds if and only if F has these properties. This is true since F is known in explicit form and is given by (See e.g. [85])

$$1 - F(t) = \Phi\left(\frac{t - 2\ln\alpha}{2\sqrt{t}}\right) - \alpha\Phi\left(\frac{t + 2\ln\alpha}{2\sqrt{t}}\right),$$

where Φ is the standard Normal cumulative distribution function. Given the supermartingale Z is continuous and decreasing, its Doob-Meyer decomposition with respect to $\mathbb{F}$ is $Z_t = 1 - (1 - Z_t)$. This implies that $1 - Z_t$ is the dual predictable projection onto $\mathbb{F}$ of the increasing process $\mathbf{1}_{\{\tau\leq t\}}$. Hence the Jeulin-Yor Theorem implies that the process $\mathbf{1}_{\{\tau\leq t\}} - \int_0^{t\wedge\tau}\frac{1}{Z_s}d(1-Z)_s$ is a $\mathbb{G}$ martingale. The integral is obviously equal to $-\ln Z_{t\wedge\tau}$, and it follows that

$$\mathbf{1}_{\{\tau\leq t\}} - \ln\frac{1}{Z_{t\wedge\tau}} = \mathbf{1}_{\{\tau\leq t\}} - \ln\frac{1}{1 - F(\int_0^{t\wedge\tau}\sigma(Y_s)^2 ds)}$$

is a $\mathbb{G}$ martingale. The compensator of τ, $\ln \frac{1}{Z_{t\wedge\tau}}$, is continuous. So τ is a totally inaccessible $\mathbb{G}$ stopping time. Finally, it is clear that the existence of an intensity is equivalent to the absolute continuity of Z_t.

Remark 4 *Since $P(\tau = \infty) = 0$, continuity and monotonicity of Z implies that τ is a so-called pseudo-stopping time, see Nikeghbali and Yor [116, Theorem 1]. In this case we have that for any $\mathbb{F}$ local martingale N, the process $N_{t\wedge\tau}$ is a $\mathbb{G}$ local martingale.*

2.2.1.3 The regular conditional distributions

We compute now the conditional probabilities $Q_t(\omega, dT) = P(\tau \in dT \mid \mathcal{F}_t)(\omega)$ and provide sufficient conditions under which they admit a density. For this it suffices to compute $P(\tau > T \mid \mathcal{F}_t)$ for every t and T. This is given in the next lemma.

Lemma 43 *For every $T \geq 0$, $t \geq 0$,*

$$P(\tau > T \mid \mathcal{F}_t) = 1 - E\{F(\int_0^T \sigma(Y_s)^2 ds) \mid \mathcal{F}_t\}.$$

Proof. If $T \leq t$ the proof of Lemma 42 gives the desired result, if we redefine the functional h to be given by

$$h(f) = h(f, \omega) = \mathbf{1}_{\{\inf_{s\leq T}(f\cdot B)_s(\omega) - \frac{1}{2}\int_0^s f(u)^2 du > \ln\alpha\}}.$$

If $T > t$, note that $P(\tau > T \mid \mathcal{F}_t) = E(P(\tau > T \mid \mathcal{F}_T) \mid \mathcal{F}_t)$ and apply Lemma 42 to conclude. ■

Under certain growth conditions on σ it is possible to show, as it is done in Lemma 44, that for each t, the conditional survival probability admits a density with respect to Lebesgue measure, which implies the existence of a $\mathbb{G}$ intensity for the random time τ.

Lemma 44 *Assume that σ is at most of polynomial growth, i.e. there exist constants C and M such that for all $y \geq 0$, $\sigma(y) \leq C(1+y)^M$. Then $Q_t(\omega, dT) = \alpha_t^T(\omega)dT$ a.s. where*

$$\alpha_t^T(\omega) = E\{F'(\int_0^T \sigma(Y_s)^2 ds)\sigma(Y_T)^2 \mid \mathcal{F}_t\}(\omega).$$

It readily follows that $1_{\{\tau\leq\cdot\}} - \int_0^{\cdot\wedge\tau} \lambda_s ds$ is a $\mathbb{G}$ martingale where

$$\lambda_t = \sigma^2(Y_t)\frac{F'(\int_0^t \sigma^2(Y_s)ds)}{1 - F(\int_0^t \sigma^2(Y_s)ds)}$$

Proof. For $T \leq t$, the result follows directly by differentiating with respect to T.

For $T > t$, we need to interchange differentiation with respect to T and expectation. First, the existence of $F'(\int_0^T \sigma^2(Y_s)ds)\sigma^2(Y_T)$ is straightforward. Let $\varepsilon > 0$ fixed and define

$I = [T-\varepsilon, T+\varepsilon]$. We use a local version of the theorem of inversion of differentation and expectation to prove that $E(F(\int_0^T \sigma^2(Y_s)ds)|\mathcal{F}_t)(\omega)$ is differentiable at the fixed time T, with the derivative given in the formula above. To do so, it suffices to prove that

$$E(\sup_{t\in I}|F^{'}(\int_0^t \sigma^2(Y_s)ds)\sigma^2(Y_t)|) < \infty$$

First, using the expression of $1-F(t)$ we obtain

$$-F^{'}(t) = \frac{t+2\ln(\alpha)}{4t^{\frac{3}{2}}}\Phi^{'}(\frac{t-2\ln(\alpha)}{2\sqrt{t}}) - \alpha\frac{t-2\ln(\alpha)}{4t^{\frac{3}{2}}}\Phi^{'}(\frac{t+2\ln(\alpha)}{2\sqrt{t}})$$

where $\Phi^{'}(t) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}$ and $\ln(\alpha)$ non zero. A straightforward study of this function gives that $\lim_{t\to\infty} -F^{'}(t) = 0$ and $-F^{'}(t)$ is equivalent in 0 to $\frac{(1+\alpha)ln(\alpha)}{2\sqrt{2\pi}t^{\frac{3}{2}}}e^{-\frac{1}{2}(\frac{ln(\alpha)}{\sqrt{t}})^2}$ which converges to 0 when t tends to 0. Since $F^{'}$ is continuous, it is bounded on $[0,\infty[$ and reaches its bounds. Therefore, there exists a constant c such that

$$\begin{aligned} E(\sup_{t\in I}|F^{'}(\textstyle\int_0^t \sigma^2(Y_s)ds)\sigma^2(Y_t)|) &\le cE(\sup_{t\in I}\sigma^2(Y_t)) \\ &\le cE(\sup_{t\le T+\varepsilon}\sigma^2(Y_t)) \\ &\le \tilde{C}(1+E(\sup_{t\le T+\varepsilon}Y_t^{2M})) \end{aligned}$$

where the last inequality follows from the assumption that σ has at most a polynomial growth and $\tilde{C} = \tilde{C}(C, M, c)$ is a constant.

It remains to prove that the expectation in the right side is finite. For this, we use Doob's maximal inequality to the positive submartingale Y_t^{2M} and obtain

$$E(\sup_{t\le T+\varepsilon} Y_t^{2M}) = \int_{\mathbb{R}^+} P(\sup_{t<T+\varepsilon} Y_t^{2M} \ge x)dx \le 1 + \int_1^\infty \frac{Y_{T+\varepsilon}^{4M}}{x^2} = 1 + E(Y_{T+\varepsilon}^{4M}) = C_I$$

Since Y is log normal, $E(Y_s^\beta) = e^{\frac{1}{2}\beta(\beta-1)s}$ for all $s \ge 0$, hence $C_I = 1 + e^{2M(4M-1)(T+\varepsilon)}$. This ends the proof. ■

Remark 5 *Note that the estimate $\Phi'(x) \le (2\pi)^{-1/2}$ is not tight. By using a more precise bound, weaker conditions on σ may be achievable. Moreover, the argument relies heavily on the fact that Y is geometric Brownian motion rather than standard Brownian motion. Note also that Lemma 44 shows that τ is a so-called initial time, as defined by Jeanblanc and Le Cam [85]. Moreover, since the process $(\alpha_t^T)_{t\ge0}$ is constant after T, [85, Corollary 1] implies that Hypothesis (H) holds between $\mathbb{F}$ and $\mathbb{G}$, i.e. every $\mathbb{F}$ martingale remains a $\mathbb{G}$ martingale.*

We can now introduce a simple multiple firms model which would be the building block of our next credit contagion model.

2.2.2 Multiple firms : A conditional independence model

We consider now the following multiple firms model.

$$\begin{aligned} dX_t^i &= X_t^i \sigma_i(Y_t) dB_t^i, \qquad i = 1, \dots, n \\ dY_t &= Y_t dW_t, \end{aligned}$$

with $\{W, B^1, \dots, B^n\}$ independent Brownian motions, $\mathbb{H}$ is their natural filtration and $\mathbb{F}$ is the natural filtration of W. We assume that the default times are defined as the first hitting times of given barriers $\alpha_i \in (0, X_0^i)$, i.e.

$$\tau_i = \inf\{t \geq 0 : X_t^i \leq \alpha_i\}$$

Let $\mathbb{G}$ be the progressive expansion of $\mathbb{F}$ with $\{\tau_1, \dots \tau_n\}$. That is,

$$\mathcal{G}_t = \bigcap_{u>t} \Big(\mathcal{F}_u \vee \bigvee_{i=1}^{n} \sigma(\tau_i \wedge s : s \leq u) \Big), \qquad \text{for each } t \geq 0.$$

The goal of this section is to compute the $\mathbb{G}$ compensators of the different default times. This can be done easily since the firm processes are independent conditionally to $\mathcal{F}_t$. The following lemma makes this more precise.

Lemma 45 *Conditionally on $\sigma(Y_s : s \leq t)$, $\{X_s^i : s \geq 0\}$ and $\{X_s^j : s \leq t, j \neq i\}$ are independent.*

Proof. Fix $0 \leq t < \infty$ and choose deterministic times $0 \leq t_1, \dots, t_m \leq t$ and $t \leq T_1, \dots, T_M < \infty$, as well as bounded real-valued Borel functions g_{jk} $(j = 1, \dots, n, k = 1, \dots, m)$ and G_k $(k = 1, \dots, M)$. We also define

$$\tilde{g}_j = \prod_{k=1,\dots,m} g_{jk}(X_{t_k}^j) \qquad \text{and} \qquad \tilde{G}_i = \prod_{k=1,\dots,M} G_k(X_{T_k}^i).$$

It suffices to prove that

$$E\left[\tilde{G}_i \prod_{j=1,\dots,n} \tilde{g}_j \,\middle|\, Y_s : s \leq t \right] = \prod_{i \neq j=1,\dots,n} E\left[\tilde{g}_j \mid Y_s : s \leq t\right] E\left[\tilde{g}_i \tilde{G}_i \mid Y_s : s \leq t\right].$$

A conditioning argument and the Markov property of X^i with respect to the filtration $\mathbb{H}$ imply that the left side is equal to

$$\begin{aligned} & E\left[\prod_{j=1,\dots,n} \tilde{g}_j E\left[\tilde{G}_i \,\middle|\, Y_s, X_s^j : s \leq t, j = 1, \dots n\right] \,\middle|\, Y_s : s \leq t \right] \\ &= E\left[\prod_{j=1,\dots,n} \tilde{g}_j E\left[\tilde{G}_i \,\middle|\, X_t^i\right] \,\middle|\, Y_s : s \leq t \right] \\ &= E\left[\prod_{j=1,\dots,n} \tilde{g}_j G(X_t^i) \,\middle|\, Y_s : s \leq t \right], \end{aligned}$$

where G is a bounded real-valued Borel function. Each $\tilde{g}_j$ ($j \neq i$) as well as $\tilde{g}_i\tilde{G}_i$ is a product measurable random function of the path of Y up to time t, independent accross different j. That is, there are product measurable functionals $h_j : \Omega \times L[0,t] \to \mathbb{R}$ such that $\tilde{g}_j = h_j(Y_s : s \leq t)$ ($j \neq i$) and $\tilde{g}_i\tilde{G}_i = h_i(Y_s : s \leq t)$, and such that the random variables $h_j(\cdot, f)$, $j = 1, \ldots, n$, are mutually independent for every $f \in L[0,t]$. Therefore, the previous quantity equals

$$\prod_{i \neq j = 1,\ldots,n} E\left[\tilde{g}_j \mid Y_s : s \leq t\right] E\left[\tilde{g}_i\tilde{G}_i \mid Y_s : s \leq t\right],$$

as required. ■

Now, for each $1 \leq i \leq n$, define the filtration $\mathbb{G}^{-i}$ as

$$\mathcal{G}_t^{-i} = \bigcap_{u>t} \Big(\mathcal{F}_u \vee \sigma(\tau_j \wedge s : s \leq u, j \neq i)\Big), \qquad \text{for each } t \geq 0,$$

i.e. the filtration we get by adding progresssively all the random times except τ_i. By the Jeulin-Yor Theorem for filtration expansion with a single time, it is enough to know the Doob-Meyer decomposition of the $\mathbb{G}^{-i}$ supermartingale $Z_t^i = P(\tau_i > t \mid \mathcal{G}_t^{-i})$ in order to find the $\mathbb{G}$ compensator of $\mathbf{1}_{\{\tau_i \leq \cdot\}}$. First,

$$Z_t^i = P(\tau_i > t \mid \mathcal{G}_t^{-i}) = P(\tau_i > t \mid Y_s, \tau_j \wedge s : s \leq t, j \neq i).$$

Due to Lemma 45, τ_i and $\{\tau_j \wedge s : s \leq t, j \neq i\}$ are conditionally independent given $\{Y_s : s \leq t\}$, so continuing the above calculation we get

$$Z_t^i = P(\tau_i > t \mid Y_s : s \leq t) = P(\tau_i > t \mid \mathcal{F}_t).$$

We know how to compute this quantity, as well as the finite variation part A^i of its Doob-Meyer decomposition: $A^i = 1 - Z^i$. The following lemma is then straigthforward.

Lemma 46 *For all $1 \leq i \leq n$, $\mathbf{1}_{\{\tau_i \leq t\}} - \ln \frac{1}{1-F(\int_0^{t\wedge\tau_i} \sigma_i(Y_s)^2 ds)}$ is a $\mathbb{G}$ martingale.*

It follows from Lemma 46 that τ_i will have an intensity if σ_i satisfies the assumption in Lemma 44. These intensities do not jump at the arrival of the other defaults and there is no information induced credit contagion effect in this model. However, by altering slightly the volatility structure in the model above, one can introduce this credit contagion effect. This is the aim of the next subsection.

2.2.3 A first structural model with credit contagion effect

We introduce a dependence structure between two firms that are assumed to be only partially observed. Consider the following setup with two firms:

$$\begin{aligned} dX_t^1 &= X_t^1 \sigma_1(Y_t) dB_t^1 \\ dX_t^2 &= X_t^2 f(X_t^1) \sigma_2(Y_t) dB_t^2 \\ dY_t &= Y_t dW_t, \end{aligned}$$

with $\{W, B^1, B^2\}$ are independent Brownian motions, $\mathbb{H}$ is their natural filtration, $\mathbb{F}$ is the natural filtration of W. Default times are again given by: $\tau_i = \inf\{t \geq 0 : X_t^i \leq \alpha_i\}$ with $\alpha_i \in (0, X_0^i)$, for $i \in \{1, 2\}$. Assume that the market information is modelled by the filtration $\mathbb{G}$, the progressive expansion of $\mathbb{F}$ with $\{\tau_1, \tau_2\}$. That is, $\mathcal{G}_t = \bigcap_{u>t} \Big(\mathcal{F}_u \vee \bigvee_{i=1}^2 \sigma(\tau_i \wedge s : s \leq u)\Big)$, for each $t \geq 0$. In this model, the two firms do not play the same role. Firm 1, whose asset value process is denoted X^1, can be thought of as a big supplier for a small firm (whose value process is X^2) and both firms depend on some exogenous observed economy factor Y. We are interested on the default intensity of firm 2 at the default time of its main supplier, firm 1. Therefore, we need to compute the $\mathbb{G}$ compensator of τ_2.

Let $\mathbb{F}^1$ be the progressive expansion of $\mathbb{F}$ with τ_1. Introduce Z^1 the optional projection of $1_{\{\tau_1 > .\}}$ onto $\mathbb{F}$ and Z^2 the optional projection of $1_{\{\tau_2 > .\}}$ onto $\mathbb{F}^1$. Let β be a Brownian motion and introduce the functions $F_i(t) = P(\inf_{\{s \leq t\}} \beta_s - \frac{1}{2}s \leq \ln(\alpha_i))$ for $i \in \{1, 2\}$. Applying the result from the first section, we know that the optional projection of $1_{\{\tau_1 > .\}}$ onto $\mathbb{F}$ is given by

$$Z_t^1 = P(\tau_1 > t \mid \mathcal{F}_t) = 1 - F_1(\int_0^t \sigma_1^2(Y_s) ds)$$

and that $1_{\{\tau_1 \leq .\}} - \ln(\frac{1}{Z^1_{.\wedge\tau_1}})$ is an $\mathbb{F}^1$ martingale. Under mild regularity assumptions on σ_1, we proved that the compensator of $1_{\{\tau_1 \leq .\}}$ is absolutely continuous w.r.t Lebesgue measure and $1_{\{\tau_1 \leq .\}} - \int_0^{.\wedge\tau_1} \lambda_s^1 ds$ is an $\mathbb{F}^1$ martingale where $\lambda_t^1 = \sigma_1^2(Y_t) \frac{F_1'(\int_0^t \sigma^2(Y_s)ds)}{1 - F_1(\int_0^t \sigma^2(Y_s)ds)}$. We focus now on the computation of the $\mathbb{G}$ compensator of $1_{\{\tau_2 \leq .\}}$. We first state and prove the following result.

Lemma 47 *The optional projection of* $1_{\{\tau_2 > .\}}$ *onto* $\mathbb{F}^1$ *is given by*

$$Z_t^2 = 1 - E\Big(F_2\big(\int_0^t f^2(X_s^1)\sigma_2^2(Y_s)ds\big) \mid \mathcal{F}_t^1\Big)$$

Proof. We introduce the filtration $\tilde{\mathcal{F}}_t^1 = \sigma(W_s, B_s^1, s \leq t)$. Since $\mathcal{F}_t^1 \subset \tilde{\mathcal{F}}_t^1$, $Z_t^2 = P(\tau_2 > t \mid \mathcal{F}_t^1) = E(P(\tau_2 > t \mid \tilde{\mathcal{F}}_t^1) \mid \mathcal{F}_t^1)$. Let us compute $P(\tau_2 > t \mid \tilde{\mathcal{F}}_t^1)$.

$$\begin{aligned}
P(\tau_2 > t \mid \tilde{\mathcal{F}}_t^1) &= E(1_{\{\tau_2 > t\}} \mid \tilde{\mathcal{F}}_t^1) = E(1_{\{\inf_{s \leq t} X_s^2 \geq \ln(\alpha_2)\}} \mid \tilde{\mathcal{F}}_t^1) \\
&= E\Big(1_{\{\inf_{s \leq t} \int_0^s f(X_u^1)\sigma_2(Y_u)dB_u^2 - \frac{1}{2}\int_0^s f^2(X_u^1)\sigma_2^2(Y_u)du \geq \ln(\alpha_2)\}} \mid \tilde{\mathcal{F}}_t^1\Big) \\
&= 1 - F_2\big(\int_0^t f^2(X_s^1)\sigma_2^2(Y_s)ds\big)
\end{aligned}$$

where the last equality follows from the same techniques as in the proof of Lemma 42. Therefore

$$Z_t^2 = P(\tau_2 > t \mid \mathcal{F}_t^1) = 1 - E\Big(F_2\big(\int_0^t f^2(X_s^1)\sigma_2^2(Y_s)ds\big) \mid \mathcal{F}_t^1\Big)$$

■

We still need to compute the Doob-Meyer decomposition of the $\mathbb{F}^1$ supermartingale Z^2. In order to do this, we need the following lemma, available in [106]. We provide a proof for completeness.

Lemma 48 *Let $\mathbb{G}$ and $\mathbb{H}$ be two filtrations such that $\mathbb{G} \subset \mathbb{H}$. Suppose a is an $\mathbb{H}$-adapted process such that $E\{\int_0^t |a_s|ds\} < \infty$ for each $t \geq 0$. Then*

$$M_t = E\Big(\int_0^t a_s ds \mid \mathcal{G}_t\Big) - \int_0^t E(a_s \mid \mathcal{G}_s)ds$$

is a $\mathbb{G}$ martingale. Here $E(a_s \mid \mathcal{G}_s)$ denotes the $\mathbb{G}$ optional projection of a.

Proof. Let $0 \leq s < t$. Then

$$\begin{aligned} E(M_t \mid \mathcal{G}_s) &= E\Big(E(\int_0^t a_u du \mid \mathcal{G}_t) - \int_0^t E(a_u \mid \mathcal{G}_u)du \mid \mathcal{G}_s\Big) \\ &= E(\int_0^t a_u du \mid \mathcal{G}_s) - E(\int_0^t E(a_u \mid \mathcal{G}_u)du \mid \mathcal{G}_s) \\ &= E(\int_0^s a_u du \mid \mathcal{G}_s) + E(\int_s^t a_u du \mid \mathcal{G}_s) - \int_0^s E(a_u \mid \mathcal{G}_u)du - \mathbb{E}(\int_s^t \mathbb{E}(a_u \mid \mathcal{G}_u)du \mid \mathcal{G}_s). \end{aligned}$$

By Fubini's theorem we obtain $E(\int_s^t E(a_u \mid \mathcal{G}_u)du \mid \mathcal{G}_s) = \int_s^t E(a_u \mid \mathcal{G}_s)du$ and $E(\int_s^t a_u du \mid \mathcal{G}_s) = \int_s^t E(a_u \mid \mathcal{G}_s)du$. Hence

$$E(M_t \mid \mathcal{G}_s) = E\Big(\int_0^t a_u du \mid \mathcal{G}_u\Big) - \int_0^s E(a_u \mid \mathcal{G}_u)du = M_s,$$

as required. ■

Lemma 49 *Assume f and σ_2 are bounded. Then $1_{\{\tau_2 \leq .\}} - \int_0^{.\wedge\tau_2} \lambda_s^2 ds$ is a $\mathbb{G}$ martingale, where*

$$\lambda_t^2 = \frac{\sigma_2^2(Y_t)}{Z_{t-}^2} E\Big(f^2(X_t^1)F_2^{'}(\int_0^t f^2(X_s^1)\sigma_2^2(Y_s)ds) \mid \mathcal{F}_t^1\Big)$$

Proof. Let $u_t = F_2(\int_0^t f^2(X_s^1)\sigma_2^2(Y_s)ds)$ and $v_t = f^2(X_t^1)\sigma_2^2(Y_t)F_2^{'}(\int_0^t f^2(X_s^1)\sigma_2^2(Y_s)ds)$ Then $du_t = v_t dt$. Since $F_2(0) = 0$, it follows that $u_t = \int_0^t v_s ds$ and we can write $Z_t^2 = 1 - E(u_t \mid \mathcal{F}_t^1) = 1 - E(\int_0^t v_s ds \mid \mathcal{F}_t^1)$. Under the boundedness assumptions of f and σ_2, we can use Lemma 48 and it follows that

$$\tilde{M}_t := E(\int_0^t v_s ds \mid \mathcal{F}_t^1) - \int_0^t E(v_s \mid \mathcal{F}_s^1)ds$$

is an $\mathbb{F}^1$ martingale. Hence $Z_t^2 = 1-(\tilde{M}_t+\int_0^t E(v_s \mid \mathcal{F}_s^1)ds)$. Define the $\mathbb{F}^1$ martingale $M_t^2 = 1-\tilde{M}_t$. Since F_2 is increasing, v_t is a.s. nonnegative and $A_t^2 := \int_0^t E(v_s \mid \mathcal{F}_s^1)ds$ is an adapted increasing process. We obtained the Doob-Meyer decomposition of the supermartingale $Z_t^2 = M_t^2 - A_t^2$. An application of Jeulin-Yor theorem gives the compensator of $1_{\{\tau_2 \leq \cdot\}}$ and

$$1_{\{\tau_2 \leq t\}} - \int_0^{t\wedge\tau_2} \frac{dA_s^2}{Z_{s-}^2}$$

is a $\mathbb{G}$ martingale. The result in the lemma follows now from the absolute continuity of A^2 and the definition of v. ■

Note that τ_2 has an absolutely continuous $\mathbb{G}$ compensator and is a totally inaccessible $\mathbb{G}$ stopping time. In order to prove the existence of a credit contagion effect, it remains to check if the expectation

$$E\Big(f^2(X_t^1)F_2'\big(\int_0^t f^2(X_s^1)\sigma_2^2(Y_s)ds\big) \mid \mathcal{F}_t^1\Big)$$

jumps at τ_1. The size of this jump would quantify the credit contagion effect. Let $U_t = f^2(X_t^1)F_2'\big(\int_0^t f^2(X_s^1)\sigma_2^2(Y_s)ds\big)$. Then $\lambda_t^2 = \frac{\sigma_2^2(Y_t)}{Z_{t-}^2}E(U_t \mid \mathcal{F}_t^1)$. Since $t \to \frac{\sigma_2^2(Y_t)}{Z_{t-}^2}$ is left continuous, λ^2 jumps at τ_1 if and only if $({}^{o\mathbb{F}^1}U)_t := E(U_t \mid \mathcal{F}_t^1)$ jumps at τ_1. The quantity of interest is then $\delta := ({}^{o\mathbb{F}^1}U)_{\tau_1} - \limsup_{t\to\tau_1^-} Q_t$, where $Q_t := 1_{\{\tau_1>t\}}E(U_t \mid \mathcal{F}_t^1)$. First,

$$({}^{o\mathbb{F}^1}U)_{\tau_1} = E(U_{\tau_1} \mid \mathcal{F}_{\tau_1}^1) = E\Big(f^2(X_{\tau_1}^1)F_2'\big(\int_0^{\tau_1} f^2(X_s^1)\sigma_2^2(Y_s)ds\big) \mid \mathcal{F}_{\tau_1}^1\Big)$$

and by a basic result from progressive filtrations expansion theory,

$$Q_t = \frac{1_{\{\tau_1>t\}}}{Z_t^1}E(1_{\{\tau_1>t\}}U_t \mid \mathcal{F}_t) = \frac{1_{\{\tau_1>t\}}}{Z_t^1}E\Big(1_{\{\tau_1>t\}}f^2(X_t^1)F_2'\big(\int_0^t f^2(X_s^1)\sigma_2^2(Y_s)ds\big) \mid \mathcal{F}_t\Big)$$

It is not obvious to prove that $\delta \neq 0$ a.s. in the general setting, so we provide a simple example where explicit computations can be done.

Example 5 *Assume $f(x) = c_1$ for all $x \in [0, \alpha_1]$ and $f(x) = c_2$, for all $x > \alpha_1$, where c_1 and c_2 are distinct positive constants. Then*

$$({}^{o\mathbb{F}^1}U)_{\tau_1} = E\Big(f^2(X_{\tau_1}^1)F_2'\big(\int_0^{\tau_1} f^2(X_s^1)\sigma_2^2(Y_s)ds\big) \mid \mathcal{F}_{\tau_1}^1\Big) = c_1^2F_2'\big(c_2^2\int_0^{\tau_1}\sigma_2^2(Y_s)ds\big)$$

and

$$\begin{aligned}
1_{\{\tau_1>t\}}E(U_t \mid \mathcal{F}_t^1) &= \frac{1_{\{\tau_1>t\}}}{Z_t^1}E\Big(1_{\{\tau_1>t\}}f^2(X_t^1)F_2'\big(\int_0^t f^2(X_s^1)\sigma_2^2(Y_s)ds\big) \mid \mathcal{F}_t\Big)\\
&= \frac{1_{\{\tau_1>t\}}}{Z_t^1}E\Big(1_{\{\tau_1>t\}}c_2^2F_2'\big(c_2^2\int_0^t \sigma_2^2(Y_s)ds\big) \mid \mathcal{F}_t\Big)\\
&= \frac{1_{\{\tau_1>t\}}}{Z_t^1}E(1_{\{\tau_1>t\}} \mid \mathcal{F}_t)c_2^2F_2'\big(c_2^2\int_0^t \sigma_2^2(Y_s)ds\big)\\
&= 1_{\{\tau_1>t\}}c_2^2F_2'\big(c_2^2\int_0^t \sigma_2^2(Y_s)ds\big)
\end{aligned}$$

This quantity converges to $c_2^2F_2'(c_2^2\int_0^{\tau_1}\sigma_2^2(Y_s)ds)$ when t increases to τ_1. Hence

$$\delta = (c_1^2 - c_2^2)F_2'\big(c_2^2\int_0^{\tau_1}\sigma_2^2(Y_s)ds\big)$$

is non zero. So, when $c_1 > c_2$*, the intensity of default of firm* 2 *increases at the default time of firm* 1*, but when* $c_1 < c_2$ *this intensity of default decreases at the default time of firm* 1*. Financially speaking, in the first case, the default of firm* 1 *is bad for firm* 2*, this would be the case for example when firm* 1 *is a big supplier for firm* 2 *or firm* 1 *is the main client of firm* 2*. In the second case, the default of firm* 1 *is a blessing for firm* 2*, which would be the case for instance if the two firms are competitors.*

In the next subsection, instead of altering the volatility structure, we introduce the credit contagion effect by modeling the firm values as two co-dependent diffusions, where the dependence is through the drift terms. The computations in the model below require the use of techniques from filtering theory and rely heavily on Kusuoka's representation theorem in the filtration expanded with the default times, after a change of measure.

2.2.4 A structural credit contagion model in finite time horizon

Fix a time horizon $T > 0$. Let $(\Omega, \mathcal{H}, \mathbb{H}, P)$ a complete filtered probability space where $\mathbb{H}$ is the natural filtration of a three-dimensional standard Brownian motion (W, W^1, W^2). Consider the model

$$\begin{aligned} dY_t &= Y_t dW_t \\ dX_t^1 &= X_t^1(dW_t^1 + \mu_1(Y_t, X_t^1, X_t^2)dt) \\ dX_t^2 &= X_t^2(dW_t^2 + \mu_2(Y_t, X_t^1, X_t^2)dt) \end{aligned}$$

where μ_1 and μ_2 are bounded measurable functions. $\mathbb{F}$ is the natural filtration of Y and $\mathbb{G}$ its progressive with τ_1 and τ_2. Writing the model under the historical probability P makes sense from a credit contagion perspective, since we are interested in the physical probability that a firm defaults and the influence of this default on the other firm. The pricing-hedging issues are beyond the scope of this study. This model is a generalization of the one studied in [113], in which the author is interested in only one time, and in particular in which no credit contagion issues are discussed. The methods used there are however useful to our own study and they turned out to be generalizable to the two times case at the expense of much more involved computations. First, under an equivalent probability measure, the processes Y, X^1 and X^2 are independent.

2.2.4.1 Measure change

Thanks to the boundedness assumptions of the functions μ_1 and μ_2, we can define a new probability measure Q equivalent to P with density

$$\frac{dQ}{dP} = \mathcal{E}\Big(\int_0^{\cdot} -\mu_1(Y_s, X_s^1, X_s^2)dW_s^1 - \mu_2(Y_s, X_s^1, X_s^2)dW_s^2\Big)_T$$

By Girsanov's Theorem, $\tilde{W}_t^1 := W_t^1 + \int_0^t \mu_1(Y_s, X_s^1, X_s^2)ds$, $\tilde{W}_t^2 := W_t^2 + \int_0^t \mu_2(Y_s, X_s^1, X_s^2)ds$ and W_t are independent Brownian motions under Q and the model can be written as follows

$$dY_t = Y_t dW_t$$
$$dX^1_t = X^1_t d\tilde{W}^1_t$$
$$dX^2_t = X^2_t d\tilde{W}^2_t$$

We obtain an *independence* model, easier than the one dealt with in section 2.2.2. It follows from Lemma 46 that

$$1_{\{\tau_1 \le t\}} - \ln\big(\frac{1}{1-F_1(t\wedge\tau_1)}\big) \qquad \text{and} \qquad 1_{\{\tau_2 \le t\}} - \ln\big(\frac{1}{1-F_2(t\wedge\tau_2)}\big)$$

are $(\mathbb{G}, Q)$ martingales. It now easy to see that τ_1 is $\mathbb{F}^1$ initial and $(\alpha^{1,u}_t)_{t\ge 0}$ is constant after u. Hence Hypothesis (H) holds between $(\mathbb{F}, Q)$ and $(\mathbb{F}^1, Q)$. Also, τ_2 is $\mathbb{F}^1$ initial and that Hypothesis (H) holds between $(\mathbb{F}^1, Q)$ and $(\mathbb{G}, Q)$. Finally, Hypothesis (H) holds between $(\mathbb{F}, Q)$ and $(\mathbb{G}, Q)$. It is also straigthforward to see that the $\mathbb{G}$ compensators of τ_1 and τ_2 are absolutely continuous with respect to Lebesgue measure i.e. there exist $\mathcal{G}_t$ adapted (in fact determinist) processes λ^1 and λ^2 such that $M^1_t := 1_{\{\tau_1\le t\}} - \int_0^{t\wedge\tau_1} \lambda^1_s ds$ and $M^2_t := 1_{\{\tau_2\le t\}} - \int_0^{t\wedge\tau_2} \lambda^2_s ds$ are $(\mathbb{G}, Q)$ martingales. Explicitly, for $i \in \{1,2\}$, $\lambda^i_s = -\frac{F'_i(s)}{1-F_i(s)}$.

2.2.4.2 $\mathbb{G}$ compensators

We have now all the ingredients to apply Kusuoka's representation theorem which we recall below. For a proof, we refer the reader to [101].

Theorem 41 (Kusuoka) *If $\mathbb{F}$ is generated by a Brownian motion W, if Hypothesis (H) holds between $\mathbb{F}$ and $\mathbb{G}$ and if each random time τ_i admits a $\mathbb{G}$ intensity λ^i, i,e. $M^i_t := 1_{\{\tau_i\le t\}} - \int_0^{t\wedge\tau_i} \lambda^i_s ds$ is a $\mathbb{G}$ martingale, then any square integrable $\mathbb{G}$ martingale L admits the martingale representation*

$$L_t = L_0 + \int_0^t \alpha_s dW_s + \int_0^t \alpha^1_s dM^1_s + \int_0^t \alpha^2_s dM^2_s$$

where α, α_1 and α_2 are $\mathbb{G}$ predictable processes.

Define $\rho := \frac{dP}{dQ}$. Then

$$\rho = \mathcal{E}\Big(\int_0^{\cdot} \mu_1(Y_s, X^1_s, X^2_s) d\tilde{W}^1_s + \mu_2(Y_s, X^1_s, X^2_s) d\tilde{W}^2_s\Big)_T$$

Introduce the Radon-Nikodym density process by setting $\rho_t := E^Q(\rho \mid \mathcal{G}_t)$. Introduce also the $\mathbb{H}$ martingale

$$f_t := E^Q(\rho \mid \mathcal{H}_t) = \mathcal{E}\Big(\int_0^{\cdot} \mu_1(Y_s, X^1_s, X^2_s) d\tilde{W}^1_s + \mu_2(Y_s, X^1_s, X^2_s) d\tilde{W}^2_s\Big)_t$$

Then $\rho_t = E^Q\Big(E^Q(\rho \mid \mathcal{H}_t) \mid \mathcal{G}_t\Big) = E^Q(f_t \mid \mathcal{G}_t)$. But Ito's formula implies

$$f_t = 1 + \int_0^t f_{s-}\mu_1(Y_s, X_s^1, X_s^2)d\tilde{W}_s^1 + \int_0^t f_{s-}\mu_2(Y_s, X_s^1, X_s^2)d\tilde{W}_s^2$$

Hence

$$\rho_t = 1 + E^Q(\int_0^t f_{s-}\mu_1(Y_s, X_s^1, X_s^2)d\tilde{W}_s^1 \mid \mathcal{G}_t) + E^Q(\int_0^t f_{s-}\mu_2(Y_s, X_s^1, X_s^2)d\tilde{W}_s^2 \mid \mathcal{G}_t) \quad (2.2)$$

Now, for any $\mathbb{H}$ predictable processes a^1 and a^2 such that $E^Q(\int_0^t |a_s^1|^2 + |a_s^2|^2 ds) < \infty$, $E^Q(\int_0^t a_s^1 d\tilde{W}_s^1 \mid \mathcal{G}_t)$ and $E^Q(\int_0^t a_s^2 d\tilde{W}_s^2 \mid \mathcal{G}_t)$ are square integrable $\mathbb{G}$ martingales. This follows from the $(\mathbb{H}, Q)$ martingale property of $\int_0^t a_s^i d\tilde{W}_s^i$, for $i \in \{1, 2\}$, and the equalities

$$E^Q(\int_0^t a_s^i d\tilde{W}_s^i \mid \mathcal{G}_t) = E^Q(E^Q(\int_0^T a_s^i d\tilde{W}_s^i \mid \mathcal{H}_t) \mid \mathcal{G}_t) = E^Q(\int_0^T a_s^i d\tilde{W}_s^i \mid \mathcal{G}_t)$$

Hence ρ_t is a square integrable $\mathbb{G}$ martingale. Thanks to Theorem 41, it admits the martingale representation

$$\rho_t = 1 + \int_0^t \alpha_s dW_s + \int_0^t \alpha_s^1 dM_s^1 + \int_0^t \alpha_s^2 dM_s^2$$

where α, α_1 and α_2 are $\mathbb{G}$ predictable processes. Since ρ_t is a.s. strictly positive, the following representation also holds

$$\rho_t = 1 + \int_0^t \rho_{s-}\beta_s dW_s + \int_0^t \rho_{s-}\beta_s^1 dM_s^1 + \int_0^t \rho_{s-}\beta_s^2 dM_s^2$$

In fact, it suffices to set $\beta_s := \frac{\alpha_s}{\rho_{s-}}$ and $\beta_s^i := \frac{\alpha_s^i}{\rho_{s-}}$ for $i \in \{1, 2\}$. We can now state the following result which gives the $(\mathbb{G}, P)$ intensity of the default times τ_1 and τ_2. The proof of this theorem follows the same lines as that of the main result in [113].

Theorem 42 *The processes* $1_{\{\tau_1 \le t\}} - \int_0^{t\wedge\tau_1} \lambda_s^1(1 + \beta_s^1)ds$ *and* $1_{\{\tau_2 \le t\}} - \int_0^{t\wedge\tau_2} \lambda_s^2(1 + \beta_s^2)ds$ *are* $(\mathbb{G}, P)$ *martingales.*

Proof. This is immediate and follows from Girsanov's Theorem. Alternatively, we prove *à la main* that $\tilde{M}_t^1 := 1_{\{\tau_1 \le t\}} - \int_0^{t\wedge\tau_1} \lambda_s^1(1 + \beta_s^1)ds$ is a $(\mathbb{G}, P)$ martingale. First note that ρ_t and M_t^1 are $(\mathbb{G}, Q)$ martingales. Hence

$$\rho_t M_t^1 - [\rho, M^1]_t = \rho_t M_t^1 - \int_0^t 1_{\{\tau_1 > s\}}\lambda_s^1 \rho_{s-}\beta_s^1 ds$$

is a $(\mathbb{G}, Q)$ martingale. We can compute the conditional expectation $E^P(\tilde{M}^1_t \mid \mathcal{G}_s)$.

$$
\begin{aligned}
E^P(\tilde{M}^1_t \mid \mathcal{G}_s) &= E^P(M^1_t - \int_0^t 1_{\{\tau_1 > u\}} \lambda^1_u \beta^1_u du \mid \mathcal{G}_s) = \frac{E^Q(\rho(M^1_t - \int_0^t 1_{\{\tau_1>u\}} \lambda^1_u \beta^1_u du) \mid \mathcal{G}_s)}{\rho_s} \\
&= \frac{E^Q(\rho_t M^1_t \mid \mathcal{G}_s)}{\rho_s} - \frac{1}{\rho_s}\Big(\int_s^t E^Q(\rho_u \lambda^1_u \beta^1_u 1_{\{\tau_1>u\}} \mid \mathcal{G}_s) du + \rho_s \int_0^s 1_{\{\tau_1>u\}} \lambda^1_u \beta^1_u \Big) \\
&= \frac{E^Q(\rho_t M^1_t - \int_s^t \rho_u \lambda^1_u \beta^1_u 1_{\{\tau_1>u\}} du \mid \mathcal{G}_s)}{\rho_s} - \int_0^s 1_{\{\tau_1>u\}} \lambda^1_u \beta^1_u du \\
&= M^1_s - \int_0^s 1_{\{\tau_1>u\}} \lambda^1_u \beta^1_u du = \tilde{M}^1_s.
\end{aligned}
$$

■

The information induced credit contagion effect will appear from the expression above as soon as there is a jump of the intensity of default of a firm at the default time of the other firm. So, the only remaining part is to determine explicitly β, β^1 and β^2. In order to do this, we need to find the martingale representations of the two $\mathbb{G}$ martingales $E^Q(\int_0^t a^1_s d\tilde{W}^1_s \mid \mathcal{G}_t)$ and $E^Q(\int_0^t a^2_s d\tilde{W}^2_s \mid \mathcal{G}_t)$. Since the proofs are similar in both cases, we prove the martingale representation for $E^Q(\int_0^t a^1_s d\tilde{W}^1_s \mid \mathcal{G}_t)$ only.
Let a be an $\mathbb{H}$ predictable process such that $E(\int_0^T |a_s|^2 ds) < \infty$. Define

$$
N = \int_0^T a_s d\tilde{W}^1_s \qquad \text{and} \qquad N_t = \int_0^t a_s d\tilde{W}^1_s \qquad \text{for each} \quad 0 \leq t \leq T
$$

Finally, consider $L_t = E^Q(\int_0^t a_s d\tilde{W}^1_s \mid \mathcal{G}_t)$. Recall that $\mathbb{G}$ is the progressive expansion of $\mathbb{F}$ with $\{\tau_1, \tau_2\}$. Clearly $L_t = E^Q(N_t \mid \mathcal{G}_t) = E^Q(N \mid \mathcal{G}_t)$.

2.2.4.3 Explicit martingale representation of L in $(\mathbb{G}, Q)$

For the rest of this subsection, all expectations are computed under the probability measure Q, so we omit it from now on. Recall that we have the $\mathbb{G}$ martingale representation

$$
L_t = \int_0^t \alpha_s dW_s + \int_0^t \alpha^1_s dM^1_s + \int_0^t \alpha^2_s dM^2_s
$$

from Theorem 41. Our goal is to give a closed form for the processes α, α^1 and α^2. Introduce the processes $N^1_t = 1_{\{\tau_1 \leq t\}}$ and $N^2_t = 1_{\{\tau_2 \leq t\}}$. Let

$$
K_t = \int_0^t c_s dW_s + \int_0^t c^1_s dM^1_s + \int_0^t c^2_s dM^2_s
$$

where c, c^1 and c^2 are bounded $\mathbb{G}$ predictable processes. The following result generalizes in a straightforward way a result in [113] to two default times.

Lemma 50 *The following identities hold*

$$E(L_tK_t) = \int_0^t E\Big(\alpha_s c_s + \alpha^1_s c^1_s \lambda^1_s(1-N^1_{s^-})ds + \alpha^2_s c^2_s \lambda^2_s(1-N^2_{s^-})\Big)ds$$
$$E(L_tK_t) = E(\int_0^t N_s c^1_s dM^1_s) + E(\int_0^t N_s c^2_s dM^2_s)$$

Proof. The first identity follows from the $\mathbb{G}$ martingale property of $L_tK_t - [L,K]_t$. First, $[W, M^i] = 0$ for $i \in \{1,2\}$. Also $[M^1, M^2] = 0$ a.s. since X^1 and X^2 cannot jump simultaneously. Hence

$$\begin{aligned} E(L_tK_t) = E([L,K]_t) &= E(\int_0^t \alpha_s c_s ds) + E(\int_0^t \alpha^1_s c^1_s dN^1_s) + E(\int_0^t \alpha^2_s c^2_s dN^2_s) \\ &= E(\int_0^t \alpha_s c_s + \alpha^1_s c^1_s \lambda^1_s(1-N^1_{s^-})ds + \alpha^2_s c^2_s \lambda^2_s(1-N^2_{s^-})ds) \\ &= \int_0^t E(\alpha_s c_s + \alpha^1_s c^1_s \lambda^1_s(1-N^1_{s^-})ds + \alpha^2_s c^2_s \lambda^2_s(1-N^2_{s^-}))ds \end{aligned}$$

For the second identity, write

$$\begin{aligned} E(L_tK_t) &= E(K_t E(\int_0^t a_s d\tilde{W}^1_s \mid \mathcal{G}_t)) = E(K_t \int_0^t a_s d\tilde{W}^1_s) \\ &= E((\int_0^t c_s dW_s + \int_0^t c^1_s dM^1_s + \int_0^t c^2_s dM^2_s)(\int_0^t a_s d\tilde{W}^1_s)) \\ &= E(\overline{M}^1_t N_t) + E(\overline{M}^2_t N_t) \end{aligned}$$

where $\overline{M}^1_t = \int_0^t c^1_s dM^1_s$ and $\overline{M}^2_t = \int_0^t c^2_s dM^2_s$. The last equality follows from the independence between W and $\tilde{W}^1$. Also, for $i \in \{1,2\}$,

$$d(\overline{M}^i_t N_t) = \overline{M}^i_{t^-} dN_t + N_t d\overline{M}^i_t + d[\overline{M}^i, N]_t$$

But $[\overline{M}^i, N]$ is null a.s. Hence $E(\overline{M}^i_t N_t) = E(\int_0^t \overline{M}^i_{s^-} dN_s) + E(\int_0^t N_s d\overline{M}^i_s)$, and since $\int_0^t \overline{M}^i_{s^-} dN_s = \int_0^t \overline{M}^i_{s^-} a_s d\tilde{W}^1_s$ is an $\mathbb{H}$ martingale, it follows that $E(\int_0^t \overline{M}^i_{s^-} dN_s) = 0$. Finally, $E(\overline{M}^i_t N_t) = E(\int_0^t N_s c^i_s dM^i_s)$ for $i \in \{1,2\}$. This proves the second result. ∎
We focus now on the computation of the terms $E(\int_0^t N_s c^i_s dM^i_s)$, $i \in \{1,2\}$. We mainly extend the approach in [113] to the two times case, the computations being then much more involved. The two expectations can be handled in the same way, and to fix ideas, we focus on the case $i = 1$. We compute $E(\int_0^t N_s c^1_s dM^1_s)$. From the definition of M^1, it follows that

$$E(\int_0^t N_s c^1_s dM^1_s) = E(\int_0^t N_s c^1_s dN^1_s) - E(\int_0^t N_s c^1_s \lambda^1_s (1-N^1_{s^-})ds)$$

The following lemma deals with the second term of the right side. Recall that $N = \int_0^T a_s d\tilde{W}^1_s$ and $N_t = \int_0^t a_s d\tilde{W}^1_s$, for each $0 \le t \le T$. Let ${}^pN$ be the predictable projection of N_t onto $\mathbb{G}$. The following holds.

Lemma 51

$$E(\int_0^t N_s c_s^1 \lambda_s^1 (1-N_{s^-}^1) ds) = \int_0^t E\Big(c_s^1 \lambda_s^1 (1-N_{s^-}^1)\ {}^p N_s\Big) ds$$
$$E(\int_0^t N_s c_s^2 \lambda_s^2 (1-N_{s^-}^2) ds) = \int_0^t E\Big(c_s^2 \lambda_s^2 (1-N_{s^-}^2)\ {}^p N_s\Big) ds$$

Proof. This holds by definition of the predictable projection. ■

Remark 6 *The predictable projection can be replaced by the predictable process* L_{s^-}*.*

We focus now on the term $E := E(\int_0^t N_s c_s^1 dN_s^1)$. Define the $\mathbb{F}$ adapted processes $K_s^1 = E((1-N_{s^-}^2)N_s \mid \mathcal{F}_s, \tau_1 = s), 0 \leq s \leq T$ and $L^1(s_1, s_2) = E(N_{s_1} \mid \mathcal{F}_{s_1}, \tau_1 = s_1, \tau_2 = s_2), 0 \leq s_2 \leq s_1 \leq T$.

Lemma 52

$$E(\int_0^t N_s c_s^1 dN_s^1) = \int_0^t E\Big(c_s^1 \lambda_s^1 (1-N_{s^-}^1)\big((1-N_{s^-}^2)\frac{K_s^1}{P(\tau_2 \geq s)} + N_{s^-}^2 L^1(s, \tau_2)\big)\Big) ds$$

Proof. We use a Monotone Class Theorem to compute E. Take c^1 of the form $c_s^1 = f(s)h_1(s \wedge \tau_1)h_2(s \wedge \tau_2)$ where f is $\mathcal{F}$ measurable bounded, h_1 and h_2 are Borel bounded functions. E becomes $E = e_1 + e_2$ where

$$e_1 = E(f(\tau_1)h_1(\tau_1)h_2(\tau_1)N_t^1 1_{\{\tau_1 \leq \tau_2\}} N_{\tau_1})$$
$$e_2 = E(f(\tau_1)h_1(\tau_1)h_2(\tau_2)N_t^1 1_{\{\tau_2 < \tau_1\}} N_{\tau_1})$$

Let us start with e_1.

$$\begin{aligned} e_1 &= E(E(f(\tau_1)h_1(\tau_1)h_2(\tau_1)N_t^1 1_{\{\tau_1 \leq \tau_2\}} N_{\tau_1} \mid \mathcal{F}_t)) \\ &= E\Big(\int_0^t f(s)h_1(s)h_2(s)E((1-N_{s^-}^2)N_s \mid \mathcal{F}_t, \tau_1 = s)P(\tau_1 \in ds)\Big) \end{aligned}$$

But since $W_u - W_s$ is independent from τ_1 and τ_2 and from $\mathcal{H}_s$, and hence from N_s, it follows that $E((1-N_{s^-}^2)N_s \mid \mathcal{F}_t, \tau_1 = s) = E((1-N_{s^-}^2)N_s \mid \mathcal{F}_s, \tau_1 = s)$. Therefore

$$e_1 = E\Big(\int_0^t f(s)h_1(s)h_2(s)K_s^1 P(\tau_1 \in ds)\Big) \tag{2.3}$$

Note that $K_s^1 = E((1-N_{s^-}^2)N_s \mid \mathcal{F}_s, \tau_1 = s), 0 \leq s \leq T$ is $\mathbb{F}$ adapted but we can choose an $\mathbb{F}$ predictable version of $E((1-N_{s^-}^2)N_s \mid \mathcal{F}_s, \tau_1 = s)$ (namely, its $\mathbb{F}$ predictable projection) in such a way that (2.3) still holds. Finally

$$\begin{aligned} e_1 &= E\Big(\int_0^t f(s)h_1(s)h_2(s)K_s^1 P(\tau_1 \geq s)\lambda_s^1 ds\Big) \\ &= \int_0^t \frac{\lambda_s^1}{P(\tau_2 \geq s)} E((1-N_{s^-}^1)(1-N_{s^-}^2)c_s^1 K_s^1) ds \end{aligned}$$

where the last equality follows from the independence of τ_1 and τ_2. We deal now with e_2.

$$\begin{aligned} e_2 &= E(E(f(\tau_1)h_1(\tau_1)h_2(\tau_2)N^1_t 1_{\{\tau_2<\tau_1\}}N_{\tau_1} \mid \mathcal{F}_t)) \\ &= E\Big(\int_0^t \int_0^{s_1} f(s_1)h_1(s_1)h_2(s_2)E(N_{s_1} \mid \mathcal{F}_t, \tau_1 = s_1, \tau_2 = s_2)P(\tau_1 \in ds_1, \tau_2 \in ds_2)\Big) \end{aligned}$$

Therefore

$$e_2 = E\Big(\int_0^t \int_0^{s_1} f(s_1)h_1(s_1)h_2(s_2)L^1(s_1,s_2)P(\tau_1 \in ds_1, \tau_2 \in ds_2)\Big) \tag{2.4}$$

using the same independence argument as for the study of e_1. For each s_2, $L^1(s_1,s_2)$ can be chosen $\mathbb{F}$ predictable (namely, take instead its $\mathbb{F}$ predictable projection) and (2.4) still holds. Finally, using the independence of τ_1 and τ_2, it follows that

$$\begin{aligned} e_2 &= \int_0^t f(s_1)h_1(s_1)E\Big(\int_0^{s_1} h_2(s_2)L^1(s_1,s_2)P(\tau_2 \in ds_2)\Big)P(\tau_1 \geq s_1)\lambda^1_{s_1}ds_1 \\ &= \int_0^t f(s_1)h_1(s_1)E\Big(h_2(s_1\wedge\tau_2)N^2_{s_1^-}L^1(s_1,\tau_2)\Big)P(\tau_1 \geq s_1)\lambda^1_{s_1}ds_1 \\ &= \int_0^t E(f(s_1)h_1(s_1\wedge\tau_1)(1-N^1_{s_1^-}))E\Big(h_2(s_1\wedge\tau_2)N^2_{s_1^-}L^1(s_1,\tau_2)\Big)\lambda^1_{s_1}ds_1 \\ &= \int_0^t E\Big(c^1_s\lambda^1_s(1-N^1_{s^-})N^2_{s^-}L^1(s,\tau_2)\Big)ds \end{aligned}$$

Putting the expressions obtained for e_1 and e_2 together and using a Monotone Class argument allows to conclude. ■

Define similarly $K^2_s = E((1-N^1_{s^-})N_s \mid \mathcal{F}_s, \tau_2 = s), 0 \leq s \leq T$ and $L^2(s_1,s_2) = E(N_{s_2} \mid \mathcal{F}_{s_2}, \tau_1 = s_1, \tau_2 = s_2), 0 \leq s_1 \leq s_2 \leq T$. We obtain a similar result for $E(\int_0^t N_s c^2_s dN^2_s)$.

Lemma 53

$$E(\int_0^t N_s c^2_s dN^2_s) = \int_0^t E\Big(c^2_s\lambda^2_s(1-N^2_{s^-})\big((1-N^1_{s^-})\frac{K^2_s}{P(\tau_1 \geq s)} + N^1_{s^-}L^2(\tau_1,s)\big)\Big)ds$$

We obtain now our representation theorem.

Theorem 43 *The martingale $L_t = E(\int_0^t a_s d\tilde{W}^1_s \mid \mathcal{G}_t)$ admits the representation*

$$\begin{aligned} L_t &= \int_0^t \lambda^1_s(1-N^1_{s^-})\Big((1-N^2_{s^-})\frac{K^1_s}{P(\tau_2 \geq s)} + N^2_{s^-}L^1(s,\tau_2) - {}^pN_s\Big)dM^1_s \\ &+ \int_0^t \lambda^2_s(1-N^2_{s^-})\Big((1-N^1_{s^-})\frac{K^2_s}{P(\tau_1 \geq s)} + N^1_{s^-}L^2(\tau_1,s) - {}^pN_s\Big)dM^2_s \end{aligned}$$

Proof. Let $\tilde{\alpha}^1_s$ be the integrand in the integral w.r.t M^1 and let $\tilde{\alpha}^2_s$ be the integrand in the integral w.r.t M^2 as they appear in Theorem 43. Putting together Lemmas 51, 52 and 53, wo obtain that for every bounded $\mathcal{G}$ predictable processes c, c^1 and c^2,

$$E(\int_0^t N_s c^1_s dM^1_s) = \int_0^t E\big(\tilde{\alpha}^1_s c^1_s \lambda^1_s (1 - N^1_{s-})\big) ds$$
$$E(\int_0^t N_s c^2_s dM^2_s) = \int_0^t E\big(\tilde{\alpha}^2_s c^2_s \lambda^2_s (1 - N^2_{s-})\big) ds$$

Using Lemma 50, it follows that

$$\int_0^t E(\alpha_s c_s + \alpha^1_s c^1_s \lambda^1_s (1 - N^1_{s-}) ds + \alpha^2_s c^2_s \lambda^2_s (1 - N^2_{s-})) ds$$
$$= \int_0^t E(\tilde{\alpha}^1_s c^1_s \lambda^1_s (1 - N^1_{s-}) ds + \tilde{\alpha}^2_s c^2_s \lambda^2_s (1 - N^2_{s-})) ds$$

Since the equality above holds for every bounded $\mathcal{G}$ predictable processes c, c^1 and c^2, the theorem follows. ∎

Recall now from equation 2.2 that

$$\rho_t = 1 + E^Q(\int_0^t f_{s-}\mu_1(Y_s, X^1_s, X^2_s) d\tilde{W}^1_s \mid \mathcal{G}_t) + E^Q(\int_0^t f_{s-}\mu_2(Y_s, X^1_s, X^2_s) d\tilde{W}^2_s \mid \mathcal{G}_t)$$

Therefore Theorem 43 can be used in theory to compute the processes β^1 and β^2 that appear in Theorem 42. However, this model seems to be not tractable in practice and checking if the intensities λ^1 and λ^2 jump respectively at τ_2 and τ_1 or computing explicitly the jumps sizes seem unrealistic. This is because the default intensities are expressed in terms of conditional expectations that can hardly be computed. However the form of the intensities makes it very likely that each one of them does jump at the default time of the other firm. Also, these intensities have been expressed in terms of functionals of the characteristics of the firm values. We will see that if the random times admit a conditional density, one could avoid itself that much trouble and the default intensities can be automatically expressed in terms of functional of these conditional densities. Before doing this, we provide in the next subsection a last tractable example of a structural model where not only the firm values are only partially observed, but the default barriers are random (see [57]). The particular form we choose for the default barrier of the second firm is the key assumption that makes the computations explicit and is inspired from [62].

2.2.5 Structural models with random default barriers

We present in this subsection two other models that feature a credit contagion effect. Let $(\Omega, \mathcal{H}, P)$ be a probability space where $\mathcal{H}$ is a σ-algebra of Ω representing the total information available on the market. We consider two firms whose asset values are continuous time processes $(X^1_t)_{0 \le t \le T}$ and $(X^2_t)_{0 \le t \le T}$ where T is a fixed time horizon. We model their default times as the first times they reach some thresholds L^1 and L^2 respectively. Let $\tau_i = \inf\{t \geq 0, X_t^i \leq L^i\}$, for $i = 1, 2$ be these first hitting times, with the convention that $\inf \emptyset = \infty$. We are interested in the credit contagion effect within this framework. In the two previous subsections, our approach was to assume that X^1 and X^2 follow co-dependent diffusions and the default barriers were constants. However, we are usually unable to do explicit computations and prove that the default time of one firm has an intensity that jumps at the default of the other firm. The approach we adopt here is slightly different: we introduce the contagion effect through the dependence of the threshold L^2 on the default time τ_1 of firm 1. In these models, the firms can be either unobservable or partially observable, the threshold L^1 either constant or a random variable independent from X^1 and X^2, and the threshold L^2 is chosen of the form $a + b1_{\{\tau_1 \leq T\}}$, where T is the fixed time horizon and a and b are two positive deterministic constants. Here the threshold that triggers the default of firm 2 is high if firm 1 defaults before T. This definition of L^2 is motivated by the fact that, financially speaking, the bankruptcy might occur altough the firm is still in a relatively healthy economic situation [62]. Here firm 1 would represent a major firm of a given industry. Its default would increase the probability of default of a smaller firm whose survival depends on the health of the major firm and might hence default although its economic situation is not too bad.

2.2.5.1 Totally unobserved firm valued

Assume X^1 and X^2 are not observable. The investor has access to the defaults information only. Let $\tau_1 = \inf\{t \geq 0, X_t^1 \leq L^1\}$, where L^1 is a random variable in $\mathcal{H}$, independent of X^1 and X^2. Denote by $\mathcal{H}_0$ the collection of P-null sets in $\mathcal{H}$. Let $\mathcal{F}_t^1 = \sigma(1_{\{\tau_1 \leq s\}}, s \leq t) \vee \mathcal{H}_0$ which represents the first firm's default information. Let $L^2 = a + b1_{\{\tau_1 \leq T\}}$ and $\tau_2 = \inf\{t \geq 0, X_t^2 \leq L^2\}$ and introduce the filtration $\mathbb{G}$ obtained by progressively enlarging $\mathcal{F}^1$ by the random default time τ_2, namely the right-continuous modification of $\mathcal{F}_t^1 \vee \sigma(\tau_2 \wedge t)$. For convenience we introduce, the running infimum processes $m_t^i = \inf_{s \leq t} X_s^i$, for $i = 1, 2$. Assume that L^1 has a prior density $P^{L^1}(dl)$ and that m^1 and m^2 have prior distribution functions $H^1(t, x)$ and $H^2(t, x)$ respectively. In this framework, we can compute as in [57] the $\mathbb{F}^1$ compensator of the increasing process $1_{\{\tau_1 \leq .\}}$ and prove that it admits an intensity if $H^1(t, x)$ is regular enough. Our contribution in this section is to compute the $\mathbb{G}$ compensator of $1_{\{\tau_2 \leq .\}}$ and its intensity when $H^2(t, x)$ is also regular enough and prove that this model provides a concrete example of a credit contagion effect.

We denote by $Z_t^1 = P(\tau_1 > t)$ the survival function of τ_1 and compute in Lemma 54 the compensator of τ_1, i.e. Λ^1 such that $1_{\{\tau_1 \leq .\}} - \Lambda_{t \wedge \tau_1}^1$ is an $\mathbb{F}^1$ martingale. Under mild regularity and domination assumptions, τ_1 has an $\mathbb{F}^1$ intensity. This is due to Giesecke [57], however we include the result and its proof for completeness.

Lemma 54 (Giesecke) *Assume that H^1 is continuous in t, for each fixed x. Then the*

$\mathbb{F}^1$ *compensator of the increasing process* $1_{\{\tau_1 \leq .\}}$ *is given by*

$$\Lambda_t^1 = -\ln(1 - \int_0^{\infty} H^1(t,l) P^{L^1}(dl))$$

If $H_t^1(t,x) = \frac{\partial H^1(t,x)}{\partial t}$ *exists and if there exists* g *such that for all* $x \geq 0$, $|H_t^1(t,x)| \leq g(x)$ *and* g *integrable with respect to* $P^{L^1}(dl)$, *then* τ_1 *admits an intensity given by*

$$\lambda_t^1 = \frac{\int_0^{\infty} H_t^1(t,l) P^{L^1}(dl)}{1 - \int_0^{\infty} H^1(t,l) P^{L^1}(dl)}$$

Proof. Using the independence of X^1 and L^1, it is easy to see that

$$Z_t^1 = P(\tau_1 > t) = P(m_t^1 > L^1) = \int_0^{\infty} P(m_t^1 > l) P^{L^1}(dl) = 1 - \int_0^{\infty} H^1(t,l) P^{L^1}(dl)$$

It follows from a classic theorem due to Dellacherie that the compensator of $1_{\{\tau_1 \leq .\}}$ is given by $\Lambda_{t \wedge \tau_1}^1$ where

$$\Lambda_t^1 = -\int_0^t \frac{dZ_s^1}{Z_{s-}^1}$$

The continuity in t of H^1 implies the continuity of Z^1 and that $\Lambda_t^1 = -\ln(Z_t^1) = -\ln(1 - \int_0^{\infty} H^1(t,l) P^{L^1}(dl))$.

The second point follows easily from Lebesgue's theorem under our domination assumption which proves the absolute continuity of the compensator ie. $\Lambda_t^1 = \int_0^t \lambda_s^1 ds$ where

$$\lambda_t^1 = \frac{\int_0^{\infty} H_t^1(t,l) P^{L^1}(dl)}{1 - \int_0^{\infty} H^1(t,l) P^{L^1}(dl)}$$

■

Since the compensator $\Lambda_{.\wedge \tau_1}^1$ is continuous, τ_1 is a totally inaccessible $\mathbb{F}^1$ stopping time. Let us introduce now $Z_t^2 = P(\tau_2 > t \mid \mathcal{F}_t^1)$, the optional projection onto $\mathbb{F}^1$ of $1_{\{\tau_2 > .\}}$, A^2 the dual predictable projection of $1_{\{\tau_2 \leq .\}}$ onto $\mathbb{F}^1$ and $\Lambda_t^2 = \int_0^t \frac{dA_s^2}{Z_{s-}^2}$. From Jeulin Yor theorem, we know that $1_{\{\tau_2 \leq t\}} - \Lambda_{t \wedge \tau_2}^2$ is a $\mathbb{G}$ martingale. Note that A^2 is also the finite variation part in the Doob-Meyer decomposition of Z^2. In view of this remark, we just need to compute Z^2 and its Doob-Meyer decomposition in order to find explicitly Λ^2.

Lemma 55 *The supermartingale* Z^2 *is given by*

$$Z_t^2 = \big(1 - H^2(t, a+b)\big) + \big(H^2(t, a+b) - H^2(t,a)\big) \frac{Z_T^1}{Z_t^1} 1_{\{\tau_1 > t\}}$$

Proof. First write

$$\begin{aligned} Z_t^2 &= P(\tau_2 > t \mid \mathcal{F}_t^1) = P(m_t^2 > L^2 \mid \mathcal{F}_t^1) = P(m_t^2 > a + b 1_{\{\tau_1 \leq T\}} \mid \mathcal{F}_t^1) \\ &= \mathbb{E}(1_{\{\frac{m_t^2 - a}{b} \geq 1_{\{\tau_1 \leq T\}}\}} \mid \mathcal{F}_t^1) \end{aligned}$$

Conditionning with respect to $\mathcal{F}_T^1$ gives $Z_t^2 = \mathbb{E}\big(\mathbb{E}(1_{\{\frac{m_t^2-a}{b} \geq 1_{\{\tau_1 \leq T\}}\}} \mid \mathcal{F}_T^1)\mathcal{F}_t^1\big)$ hence

$$\begin{aligned}
Z_t^2 &= \mathbb{E}(1_{\{\tau_1 \leq T\}} P(\frac{m_t^2 - a}{b} \geq 1) + 1_{\{\tau_1 > T\}} P(\frac{m_t^2 - a}{b} \geq 0) \mid \mathcal{F}_t^1) \\
&= \mathbb{E}(P(m_t^2 \geq a+b)(1 - 1_{\{\tau_1 > T\}}) + 1_{\{\tau_1 > T\}} P(m_t^2 \geq a) \mid \mathcal{F}_t^1) \\
&= P(m_t^2 \geq a+b) + \mathbb{E}(1_{\{\tau_1 > T\}}(P(m_t^2 \geq a) - P(m_t^2 > a+b)) \mid \mathcal{F}_t^1) \\
&= P(m_t^2 \geq a+b) + P(\tau_1 > T \mid \mathcal{F}_t^1) P(a \leq m_t^2 < a+b)
\end{aligned}$$

But $P(\tau_1 > T \mid \mathcal{F}_t^1) = 1_{\{\tau_1 > t\}} \frac{P(\tau_1 > T)}{P(\tau_1 > t)} = 1_{\{\tau_1 > t\}} \frac{Z_T^1}{Z_t^1}$ and $P(m_t^2 \geq c) = H^2(t,c)$ for all $c \geq 0$, hence

$$Z_t^2 = \big(1 - H^2(t, a+b)\big) + \big(H^2(t,a+b) - H^2(t,a)\big)\frac{Z_T^1}{Z_t^1} 1_{\{\tau_1 > t\}}$$

■

More explicitly, we proved that

$$Z_t^2 = P(m_t^2 \geq a+b) + P(a \leq m_t^2 < a+b) \frac{\int_0^\infty P(m_T^1 > l) P^{L^1}(dl)}{\int_0^\infty P(m_t^1 > l) P^{L^1}(dl)} 1_{\{\tau_1 > t\}}$$

In Lemma 56, we compute the Doob-Meyer decomposition of Z^2.

Lemma 56 *Assume that the increasing process $H^2(t,c) = P(m_t^2 \leq c)$ is differentiable with respect to t and define $H_t^2(s,c) = \frac{\partial H^2(s,c)}{\partial t}$ then*

$$A_t^2 = \int_0^t \Big(H_t^2(s, a+b)\big(1 - 1_{\{\tau_1 > s\}} \frac{Z_T^1}{Z_s^1}\big) + 1_{\{\tau_1 > s\}} \frac{Z_T^1}{Z_s^1} H_t^2(s,a)\Big) ds$$

Proof. Recall that $M_t^1 = 1_{\{\tau_1 \leq t\}} - \int_0^t 1_{\{\tau_1 > s\}} \lambda_s^1 ds$ is an $\mathbb{F}^1$ martingale and define $\kappa_t = (H^2(t,a+b) - H^2(t,a))\frac{Z_T^1}{Z_t^1}$ and $f_t = 1 - H^2(t,a+b) + \kappa_t(1 - \int_0^t 1_{\{\tau_1 > s\}} \lambda_s^1 ds)$. Then $Z_t^2 = f_t - \kappa_t M_t^1$.

By assumption κ_t is a continuous process with finite variation, hence $d(\kappa_t M_t^1) = M_{t^-}^1 d\kappa_t + \kappa_t dM_t^1$. Hence $dZ_t^2 = df_t - M_{t^-}^1 d\kappa_t - \kappa_t dM_t^1$. We focus now on $df_t - M_{t^-}^1 d\kappa_t$.

$$\begin{aligned}
df_t - M_{t^-}^1 d\kappa_t &= -H_t^2(t, a+b) dt \\
&+ (1 - \int_0^t 1_{\{\tau_1 > s\}} \lambda_s^1 ds) d\kappa_t - \kappa_t 1_{\{\tau_1 > t\}} \lambda_t^1 dt - (1 - 1_{\{\tau_1 > t\}} - \int_0^t 1_{\{\tau_1 > s\}} \lambda_s^1 ds) d\kappa_t \\
&= -H_t^2(t, a+b) dt + 1_{\{\tau_1 > t\}}(d\kappa_t - \kappa_t \lambda_t^1 dt)
\end{aligned}$$

By definition

$$\begin{aligned}
d\kappa_t &= Z_T^1\Big((H_t^2(t,a+b) - H_t^2(t,a))\frac{1}{Z_t^1} dt + (H^2(t,a+b) - H^2(t,a))\frac{-dZ_t^1}{(Z_t^1)^2}\Big) \\
&= Z_T^1\Big((H_t^2(t,a+b) - H_t^2(t,a))\frac{1}{Z_t^1} dt\Big) + \kappa_t \frac{-dZ_t^1}{Z_t^1}
\end{aligned}$$

and on $\{\tau_1 > t\}$, $\lambda^1_t dt = \frac{-dZ^1_t}{(Z^1_t)^2}$, hence

$$1_{\{\tau_1>t\}}(d\kappa_t - \kappa_t\lambda^1_t dt) = 1_{\{\tau_1>t\}}(H^2_t(t,a+b) - H^2_t(t,a))\frac{Z^1_T}{Z^1_t}dt$$

Hence

$$\begin{aligned}-df_t + M^1_t d\kappa_t &= -\Big(- H^2_t(t,a+b) + 1_{\{\tau_1>t\}}(H^2_t(t,a+b) - H^2_t(t,a))\frac{Z^1_T}{Z^1_t}\Big)dt \\ &= \Big(H^2_t(t,a+b)(1 - 1_{\{\tau_1>t\}}\frac{Z^1_T}{Z^1_t}) + 1_{\{\tau_1>t\}}\frac{Z^1_T}{Z^1_t}H^2_t(t,a)\Big)dt\end{aligned}$$

Since H^2 is increasing in t, $H^2_t(s,c)$ is nonnegative for all s and c, and it is clear that $Z^1_t = P(\tau_1 > t) \geq P(\tau_1 > T) = Z^1_T$, the process

$$a^2_t = \Big(H^2_t(t,a+b)(1 - 1_{\{\tau_1>t\}}\frac{Z^1_T}{Z^1_t}) + 1_{\{\tau_1>t\}}\frac{Z^1_T}{Z^1_t}H^2_t(t,a)\Big)dt$$

is non negative and $A^2_t = \int_0^t a^2_s ds$ is an increasing $\mathbb{F}^1$ predictable process. Finally $Z^2_t = \int_0^t -\kappa_s dM^1_s - A^2_t$ is the Doob-Meyer decomposition of the supermartingale Z^2. ■

Jeulin Yor theorem implies that $1_{\{\tau_2\leq t\}} - \int_0^{t\wedge\tau_2}\frac{dA^2_s}{Z^2_{s-}}$ is a $\mathbb{G}$ martingale. Since A^2 is absolutely continuous with respect to Lebesgue measure, it is also the case for the compensator Λ^2_t, and the random time τ_2 has a $\mathbb{G}$ intensity given by $\lambda^2_t = \frac{a^2_t}{Z^2_{t-}}$. It is clear from this formula that this default intensity jumps at the default time τ_1 of the first firm. This can be interpreted as information induced credit contagion, since the conditional default intensity of firm 2 jumps in response to the information that firm 1 has defaulted.

2.2.5.2 Partially observed firm values

We consider now the following model where the original filtration is non trivial, and the firm processes are partially observed. Consider three independent brownian motions W, W^1 and W^2 in $\mathcal{H}$ and define the filtration $\mathcal{F}_t = \sigma(W_s, s \leq t) \vee \mathcal{H}_0$ which represents the non default information available to the investor. We assume that the value processes satisfy

$$\begin{aligned}dY_t &= Y_t dW_t \\ dX^1_t &= X^1_t\sigma(Y_t)dW^1_t \\ dX^2_t &= X^2_t\sigma(Y_t)dW^2_t\end{aligned}$$

Consider again $\tau_1 = \inf\{t \geq 0, X^1_t \leq L^1\}$. We can assume either that L^1 is a random variable in $\mathcal{H}$ independent from $\mathcal{F}_\infty$, X^1 and X^2 with density $P^{L^1}(dl)$ with support on $(0, X^1_0)$ or that it is a deterministic constant smaller than X^1_0. We can deal with both cases in the same way. We already proved the following lemma, which we recall for clarity.

Lemma 57 *For each $l < X_0^1$, $P(m_t^1 > l \mid \mathcal{F}_t) = 1 - F_l(\int_0^t \sigma^2(Y_s)ds)$ where, for each $u > 0$,*

$$1 - F_l(u) = \Phi(\frac{u - 2\ln(l)}{2\sqrt{u}}) - l\Phi(\frac{u + 2\ln(l)}{2\sqrt{u}})$$

where Φ is the standard cumulative distribution function.

As in [57], the following holds.

Lemma 58 *The $\mathbb{F}^1$ compensator of the random time τ_1 is given by $\Lambda^1_{t\wedge\tau_1}$ where*

$$\Lambda_t^1 = -\ln\Big(\int_0^\infty \big(1 - F_l(\int_0^t \sigma^2(Y_s)ds)\big)P^{L^1}(dl)\Big)$$

This compensator is absolutely continuous with respect to Lebesgue measure and the $\mathbb{F}^1$ intensity of τ_1 is given by

$$\lambda_t^1 = \sigma^2(Y_t)\frac{\int_0^\infty F_l'(\int_0^t \sigma^2(Y_s)ds)P^{L^1}(dl)}{Z_t^1}$$

Proof. It is easy to see that

$$Z_t^1 = P(\tau_1 > t \mid \mathcal{F}_t) = P(m_t^1 > L^1 \mid \mathcal{F}_t) = \int_0^\infty P(m_t^1 > l \mid \mathcal{F}_t)P^{L^1}(dl)$$

Hence $Z_t^1 = \int_0^\infty \big(1 - F_l(\int_0^t \sigma^2(Y_s)ds)\big)P^{L^1}(dl)$. Since $t \to F_l(\int_0^t \sigma^2(Y_s)ds)$ is increasing and continuous, Z^1 is decreasing and continuous and it is a classic result due to Jeulin and Yor that

$$1_{\{\tau_1 \le t\}} - \ln(\frac{1}{Z^1_{t\wedge\tau_1}})$$

is an $\mathbb{F}^1$ martingale, which proves the first point. The $\mathbb{F}^1$ stopping time τ_1 admits an $\mathbb{F}^1$ intensity if Z^1 is absolutely continuous with respect to Lebesgue measure and that this is the case follows from regularity and good domination properties of the functions F_l. A straightforward derivation proves that this intensity is given by

$$\lambda_t^1 = \sigma^2(Y_t)\frac{\int_0^\infty F_l'(\int_0^t \sigma^2(Y_s)ds)P^{L^1}(dl)}{\int_0^\infty \big(1 - F_l(\int_0^t \sigma^2(Y_s)ds)\big)P^{L^1}(dl)} = \sigma^2(Y_t)\frac{\int_0^\infty F_l'(\int_0^t \sigma^2(Y_s)ds)P^{L^1}(dl)}{Z_t^1}$$

■

It is worth noticing that when the threshold L^1 is known, equal to a deterministic constant l_0, the previous intensity simplifies to $\lambda_t^1 = \sigma^2(Y_t)\frac{F_{l_0}'(\int_0^t \sigma^2(Y_s)ds)}{1-F_{l_0}(\int_0^t \sigma^2(Y_s)ds)} = \sigma^2(Y_t)\frac{F_{l_0}'(\int_0^t \sigma^2(Y_s)ds)}{Z_t^1}$. In both cases, Λ^1 is continuous, which is equivalent to the fact that τ_1 is a totally inaccessible $\mathbb{F}^1$ stopping time. We introduce as in the previous subsection the supermartingale $Z_t^2 = P(\tau_2 > t \mid \mathcal{F}_t^1)$, where $\tau_2 = \inf\{t \ge 0, X_t^2 \le L^2\}$ and $L^2 = a + b1_{\{\tau_1 \le T\}}$. Lemma 59 gives an explicit formula for Z^2.

Lemma 59

$$Z_t^2 = 1 - F_{a+b}(\int_0^t \sigma^2(Y_s)ds) + (F_{a+b}(\int_0^t \sigma^2(Y_s)ds) - F_a(\int_0^t \sigma^2(Y_s)ds))\frac{\mathbb{E}(Z_T^1 \mid \mathcal{F}_t)}{Z_t^1} 1_{\{\tau_1 > t\}}$$

Proof. $Z_t^2 = P(\tau_2 > t \mid \mathcal{F}_t^1) = P(m_t^2 > L^2 \mid \mathcal{F}_t^1) = P(\frac{m_t^2 - a}{b} \geq 1_{\{\tau_1 \leq T\}} \mid \mathcal{F}_t^1)$. Let $Y_t = f_t h(t \wedge \tau_1)$, where f_t is a bounded $\mathcal{F}_t$ measurable random variable and h is a measurable bounded function. Then

$$\begin{aligned}
\mathbb{E}(Y_t 1_{\{\frac{m_t^2-a}{b} \geq 1_{\{\tau_1 \leq T\}}\}}) &= \mathbb{E}(f_t h(t) 1_{\{t < \tau_1\}} 1_{\{\frac{m_t^2-a}{b} \geq 1_{\{\tau_1 \leq T\}}\}}) + \mathbb{E}(f_t h(\tau_1) 1_{\{t \geq \tau_1\}} 1_{\{\frac{m_t^2-a}{b} \geq 1_{\{\tau_1 \leq T\}}\}}) \\
&= \mathbb{E}(f_t h(t) \mathbb{E}(1_{\{t < \tau_1\}} 1_{\{\frac{m_t^2-a}{b} \geq 1_{\{\tau_1 \leq T\}}\}} \mid \mathcal{F}_t)) + \mathbb{E}(f_t \mathbb{E}(h(\tau_1) 1_{\{t \geq \tau_1\}} 1_{\{\frac{m_t^2-a}{b} \geq 1\}} \mid \mathcal{F}_t))
\end{aligned}$$

Since τ_1 and m^2 are independent conditionnally on $\mathcal{F}_t$, it readily follows that the second expectation is equal to $E_2 := \mathbb{E}(f_t \mathbb{E}(h(\tau_1) 1_{\{t \geq \tau_1\}} 1_{\{\frac{m_t^2-a}{b} \geq 1\}} \mid \mathcal{F}_t)) = \mathbb{E}(f_t \mathbb{E}(h(\tau_1) 1_{\{t \geq \tau_1\}} \mid \mathcal{F}_t) P(m_t^2 \geq a + b \mid \mathcal{F}_t)) = \mathbb{E}(f_t h(t \wedge \tau_1) 1_{\{t \geq \tau_1\}} P(m_t^2 \geq a + b \mid \mathcal{F}_t))$

We focus now on the first expectation $E_1 := \mathbb{E}(f_t h(t) \mathbb{E}(1_{\{t < \tau_1\}} 1_{\{\frac{m_t^2-a}{b} \geq 1_{\{\tau_1 \leq T\}}\}} \mid \mathcal{F}_t))$.

$$E_1 = \mathbb{E}(f_t h(t) \mathbb{E}(1_{\{t < \tau_1\}} 1_{\{T < \tau_1\}} 1_{\{\frac{m_t^2-a}{b} \geq 0\}} \mid \mathcal{F}_t)) + \mathbb{E}(f_t h(t) \mathbb{E}(1_{\{t < \tau_1\}} 1_{\{\tau_1 \leq T\}} 1_{\{\frac{m_t^2-a}{b} \geq 1\}} \mid \mathcal{F}_t))$$

Note that

$$1_{\{t < \tau_1\}} 1_{\{\tau_1 \leq T\}} = 1_{\{t < \tau_1\}} (1 - 1_{\{\tau_1 > T\}}) = 1_{\{t < \tau_1\}} - 1_{\{t < \tau_1\}} 1_{\{\tau_1 > T\}} = 1_{\{t < \tau_1\}} - 1_{\{\tau_1 > T\}}$$

Hence

$$\begin{aligned}
E_1 &= \mathbb{E}\Big(f_t h(t) \big(\mathbb{E}(1_{\{T < \tau_1\}} 1_{\{m_t^2 \geq a\}} \mid \mathcal{F}_t) + \mathbb{E}(1_{\{t < \tau_1\}} 1_{\{m_t^2 \geq a+b\}} - 1_{\{T < \tau_1\}} 1_{\{m_t^2 \geq a+b\}} \mid \mathcal{F}_t)\big)\Big) \\
&= \mathbb{E}\Big(f_t h(t) \big(\mathbb{E}(1_{\{T < \tau_1\}} 1_{\{a \leq m_t^2 < a+b\}} \mid \mathcal{F}_t) + \mathbb{E}(1_{\{t < \tau_1\}} 1_{\{m_t^2 \geq a+b\}} \mid \mathcal{F}_t)\big)\Big) \\
&= \mathbb{E}\Big(f_t h(t) \mathbb{E}(1_{\{t < \tau_1\}} \mid \mathcal{F}_t) \big(\frac{P(\tau_1 > T \mid \mathcal{F}_t) P(a \leq m_t^2 < a + b \mid \mathcal{F}_t)}{Z_t^1} + P(m_t^2 \geq a + b \mid \mathcal{F}_t)\big)\Big) \\
&= \mathbb{E}\Bigg(\mathbb{E}\Big(f_t h(t) 1_{\{\tau_1 > t\}} \big(\frac{P(\tau_1 > T \mid \mathcal{F}_t) P(a \leq m_t^2 < a + b \mid \mathcal{F}_t)}{Z_t^1} + P(m_t^2 \geq a + b \mid \mathcal{F}_t)\big) \mid \mathcal{F}_t\Big)\Bigg) \\
&= \mathbb{E}\Big(Y_t 1_{\{\tau_1 > t\}} \big(\frac{\mathbb{E}(Z_T^1 \mid \mathcal{F}_t)}{Z_t^1} P(a \leq m_t^2 < a + b \mid \mathcal{F}_t) + P(m_t^2 \geq a + b \mid \mathcal{F}_t)\big)\Big)
\end{aligned}$$

Putting E_1 and E_2 together, and using a Monotone Class Theorem, we obtain

$$Z_t^2 = P(m_t^2 \geq a + b \mid \mathcal{F}_t) + 1_{\{\tau_1 > t\}} \frac{\mathbb{E}(Z_T^1 \mid \mathcal{F}_t)}{Z_t^1} P(a \leq m_t^2 < a + b \mid \mathcal{F}_t)$$

The statement of the lemma follows since $P(m_t^2 \geq c \mid \mathcal{F}_t) = 1 - F_c(\int_0^t \sigma^2(Y_s)ds)$. ∎

We need now to compute the Doob Meyer decomposition of the supermartingale Z^2. Introduce the family of increasing processes $A_\alpha(t) = F_\alpha(\int_0^t \sigma^2(Y_s)ds)$ and the $\mathbb{F}$ martingale $N_s^1 = \mathbb{E}(Z_T^1 \mid \mathcal{F}_s)$. Then the following result holds.

Lemma 60 *The compensator of Z^2 is given by*

$$A_t^2 = \int_0^t (1 - 1_{\{\tau_1 > s\}} \frac{N_s^1}{Z_s^1}) dA_{a+b}(s) + 1_{\{\tau_1 > s\}} \frac{N_s^1}{Z_s^1} dA_a(s)$$

Proof. First recall that $M_t^1 = 1_{\{\tau_1 \le t\}} - \int_0^t \lambda_s^1 1_{\{\tau_1 > s\}} ds$ is an $\mathbb{F}^1$ martingale and that $\Lambda_t^1 = \int_0^t \lambda_s^1 ds = -\ln(Z_t^1)$. Introduce the following quantities

$$\kappa_t = P(a \le m_t^2 < a+b) \frac{N_s^1}{Z_s^1} = (A_{a+b}(t) - A_a(t)) \frac{N_s^1}{Z_s^1}$$

$$f_t = P(m_t^2 \ge a+b \mid \mathcal{F}_t) + (1 - \int_0^t \lambda_s^1 1_{\{\tau_1 > s\}} ds)\kappa_t = 1 - A_{a+b}(t) + (1 - \int_0^t \lambda_s^1 1_{\{\tau_1 > s\}} ds)\kappa_t$$

Then $Z_t^2 = f_t - \kappa_t M_t^1$ and since κ is continuous, $dZ_t^2 = -\kappa_t dM_t^1 - (M_{t-}^1 d\kappa_t - df_t)$. Straightforward computations give

$$M_t^1 d\kappa_t - df_t = (1_{\{\tau_1 \le t\}} - \int_0^t \lambda_s^1 1_{\{\tau_1 > s\}} ds) d\kappa_t + dA_{a+b}(t) - (1 - \int_0^t \lambda_s^1 1_{\{\tau_1 > s\}} ds) d\kappa_t + \lambda_t^1 1_{\{\tau_1 > t\}} \kappa_t dt$$

Introduce now the continuous process with finite variation $B_t = \frac{A_{a+b}(t) - A_a(t)}{Z_t^1}$. Then $\kappa_t = B_t N_t^1$ and $d\kappa_t = B_t dN_t^1 + N_t^1 dB_t$. Since on $\{\tau_1 > t\}$, $-\frac{dZ_t^1}{Z_t^1} = \lambda_t^1 dt$, it easy to see that

$$dB_t = \frac{dA_{a+b}(t) - dA_(t)}{Z_t^1} + B_t \lambda_t^1 dt$$

Hence $dZ_t^2 = -\kappa_t dM_t^1 + 1_{\{\tau_1 > t\}} B_t dN_t^1 - \Big(-1_{\{\tau_1 > t\}} N_t^1 \big(\frac{dA_{a+b}(t) - dA_a(t)}{Z_t^1} + B_t \lambda_t^1 dt\big) + dA_{a+b}(t) + \lambda_t^1 1_{\{\tau_1 > t\}} \kappa_t dt\Big)$. It follows that $dZ_t^2 = -\kappa_t dM_t^1 + 1_{\{\tau_1 > t\}} \frac{\kappa_t}{N_t^1} dN_t^1 - dA_t^2$ where

$$A_t^2 = \int_0^t (1 - 1_{\{\tau_1 > s\}} \frac{N_s^1}{Z_s^1}) dA_{a+b}(s) + 1_{\{\tau_1 > s\}} \frac{N_s^1}{Z_s^1} dA_a(s)$$

Since Z^1 is a supermartingale, $0 \le N_s^1 \le Z_s^1$, it is clear that $1 - 1_{\{\tau_1 > s\}} \frac{N_s^1}{Z_s^1} \ge 0$. Also for each $\alpha > 0$, A_α is an increasing continuous $\mathbb{F}$ predictable process absolutely continuous with respect to Lebesgue measure hence A^2 is thus an increasing $\mathbb{F}^1$ predictable process absolutely continuous with respect to Lebesgue measure and $1 - \int_0^t \kappa_s dM_s^1 + \int_0^t 1_{\{\tau_1 > s\}} \frac{\kappa_s}{N_s^1} dN_s^1$ is a $\mathbb{F}^1$ local martingale. By an application of Jeulin-Yor theorem, $1_{\{\tau_2 \le t\}} - \int_0^{t \wedge \tau_2} \frac{dA_s^2}{Z_{s-}^2}$ is an $\mathbb{G}$ martingale. The absolute continuity of A^2 implies that τ_2 has an $\mathbb{G}$ intensity given by $\lambda_t^2 = \frac{a_t^2}{Z_{t-}^2}$ where

$$a_t^2 = \sigma^2(Y_t)\Big(F'_{a+b}(\int_0^t \sigma^2(Y_s)ds) + 1_{\{\tau_1 > t\}} \frac{N_t^1}{Z_t^1} \big(F'_a(\int_0^t \sigma^2(Y_s)ds) - F'_{a+b}(\int_0^t \sigma^2(Y_s)ds)\big)\Big)$$

■

Again, we notice that the intensity λ^2 jumps at the default time of firm 1, which can be interpreted as an information induced credit contagion effect, since the default intensity of firm 2 jumps in response to the information that firm 1 has defaulted. To summarize, we provided in this section numerous examples of structural models that either clearly (or are likely to) exhibit an information induced credit contagion effect. We hope the computations involved convinced the reader that such models tend to be quickly non tractable. However, many structural models will be such that the default times have a conditional density w.r.t the base filtration $\mathbb{F}$. In this case, the approach to compute the default intensities can be unified, since these intensities can be expressed only in terms of the conditional densities of the random times. In the next section, we reconcile the structural approach and the reduced form approach for these types of models, and among other things (for instance analyzing the impact of ordering the default times before expanding the filtration, from a risk management perspective), we will quantify the credit contagion effect under this conditional density assumption.

2.3 Credit contagion under the conditional density assumption

In this section, the classical reduced-form and filtration expansion framework in credit risk is extended to the case of multiple, non-ordered defaults, assuming that conditional densities of the default times exist. The results in subsections 2.3.1 and 2.3.2 have been established by El Karoui, Jeanblanc and Jiao in [40] and [41]. We use their methodology to extend a bit their results, from the ranked to the non-ranked case. The extension follows directly from the results in [40] and [41], but it involves slightly complicated combinatorial techniques, hence we include proofs. Intensities and pricing formulas are derived, revealing how information-driven default contagion arises in these models. The reconciliation between structural and reduced form models is achieved for structural models where the default times admit a conditional density: in these models, the conditional densities will be functionals of the characteristics of the firm values, therefore the intensities, which are proven to exist, are also expressed in terms of the parameters of the firm values. We then analyze the impact of ordering the default times before expanding the filtration. While not important for pricing, the effect is significant in the context of risk management, and becomes even more pronounced for highly correlated and asymmetrically distributed defaults. We provide a general scheme for constructing and simulating the default times, given that a model for the conditional densities has been chosen. We then give few conditional density models and study the information induced credit contagion effect within them. Finally, we propose a toy example of a multiple firm structural model where the conditional densities are computable, in order to give an explicit example of such reconciliation. Throughout, we assume that a probability space $(\Omega, \mathcal{F}, P)$ is given where all random variables, processes, etc. are defined. We also take as given a filtration $\mathbb{F}$, satisfying the usual conditions of P completeness and right-continuity.

2.3.1 The conditional density assumption : the case of two non ranked random times

The notational complexity in the general case of arbitrarily many non-ordered default times turns out to be quite significant. For this reason we will devote the present section to the corresponding results for the special case of two times, still non-ordered. The goal is to give a flavor of the main conclusions, and to facilitate discussion of contagion effects as well as a comparison with the case of ranked times. The majority of the results stated here are corollaries of lemmas and theorems in subsequent sections, in which cases we refrain from giving separate proofs.

For the remainder of this section, we consider two default times τ_1 and τ_2, and assume the existence of $\mathcal{F}_t$-conditional densities $p_t(u_1, u_2)$:

$$p_t(u_1, u_2)du_1du_2 = P(\tau_1 \in du_1,\ \tau_2 \in du_2 \mid \mathcal{F}_t).$$

The market filtration $\mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$ is the progressive expansion of $\mathbb{F}$ with (τ_1, τ_2), namely

$$\mathcal{G}_t = \bigcap_{u>t} \mathcal{F}_u \vee \sigma(\tau_1 \wedge u,\ \tau_2 \wedge u).$$

We first write down the $\mathbb{G}$ intensities of τ_1 and τ_2. This is an immediate corollary of Theorem 45 of Section 2.3.2.

Corollary 14 *For $k = 1, 2$ the processes*

$$\mathbf{1}_{\{\tau_k \leq t\}} - \int_0^{t\wedge\tau_k} \lambda_s^k ds$$

are $\mathbb{G}$ martingales, where

$$\lambda_t^1 = \mathbf{1}_{\{t<\tau_2\}} \frac{\int_t^\infty p_t(t, u_2)du_2}{\int_t^\infty \int_t^\infty p_t(u_1, u_2)du_2du_1} + \mathbf{1}_{\{t\geq\tau_2\}} \frac{p_t(t, \tau_2)}{\int_t^\infty p_t(u_1, \tau_2)du_1}$$

and

$$\lambda_t^2 = \mathbf{1}_{\{t<\tau_1\}} \frac{\int_t^\infty p_t(u_1, t)du_1}{\int_t^\infty \int_t^\infty p_t(u_1, u_2)du_1du_2} + \mathbf{1}_{\{t\geq\tau_1\}} \frac{p_t(\tau_1, t)}{\int_t^\infty p_t(\tau_1, u_2)du_2}.$$

Alternatively, we may write this using conditional survival probabilities, rather than densities. Define

$$P_t(u_1, u_2) = P(\tau_1 > u_1,\ \tau_2 > u_2 \mid \mathcal{F}_t) = \int_{u_1}^\infty \int_{u_2}^\infty p_t(v_1, v_2)dv_2dv_1.$$

Then

$$\lambda_t^1 = -\mathbf{1}_{\{t<\tau_2\}} \frac{\partial_{u_1} P_t(t, t)}{P_t(t, t)} - \mathbf{1}_{\{t\geq\tau_2\}} \frac{\partial_{u_1,u_2} P_t(t, \tau_2)}{\partial_{u_2} P_t(t, \tau_2)},$$

where ∂_{u_i} denotes partial derivative with respect to the i:th argument and ∂_{u_1,u_2} denotes partial derivative with respect to both arguments.

The form of the $\mathbb{G}$ intensities makes it clear that they jump on the arrival of the other default, unless the conditional densities have some special structure (one example is conditional independence—see Section 2.3.2.) This is a partial result toward the reconciliation of strutural and reduced form models, at least when the default times in the structural model admit a conditional density. The reason is that the conditional densities of the default times will be a functional of the parameters of the firm values, and therefore our formulas above express the intensities in terms of these parameters of the firm values. That a given intensity jumps at the default of the other firm is then very likely, unless again the model has some very particular structure. As demonstrated through examples in the previous section, it is however usually hard to prove *à la main* that such intensities jump, see for instance subsection 2.2.4. However, in several models, the computations can be worked out explicitly, see for instance Example 5 and the example of subsection 2.2.5 where the credit contagion effect can be quantified. We will propose another toy example (see subsection 2.3.6), where we will directly apply the formulas of Corollary 14 to a two firms structural model where the default times admit a conditional density.

Knowledge of the intensities allows also one to price defaultable bonds on each of the two firms. To price more complicated products, more work is needed. As an example, and for later use, we will give a pricing formula for a security that pays one unit of currency at time T if at most one default happens before T. That is, if we define

$$\sigma_1 = \tau_1 \wedge \tau_2 \qquad \text{and} \qquad \sigma_2 = \tau_1 \vee \tau_2,$$

the time T payoff is $\mathbf{1}_{\{\sigma_2 > T\}}$. This forms a component of the cash flow to a simple synthetic CDO, or a second-to-default swap, for example. In order to focus on the issue at hand, which is the impact of the ordering of the defaults, we make the simplifying assumption of zero (or deterministic) interest rates. Assuming that the probability measure P can be used as pricing measure, the following corollary of Lemma 67 gives the price at time t.

Corollary 15 *Given the notation and assumptions of this section, we have*

$$\begin{aligned} P(\sigma_2 > T \mid \mathcal{G}_t) = {} & \mathbf{1}_{\{\tau_1 > t\} \cap \{\tau_2 > t\}} \frac{P_t(T,t) + P_t(t,T) - P_t(T,T)}{P_t(t,t)} \\ & + \mathbf{1}_{\{\tau_1 > t\} \cap \{\tau_2 \leq t\}} \frac{\partial_{u_2} P_t(T,\tau_2)}{\partial_{u_2} P_t(t,\tau_2)} + \mathbf{1}_{\{\tau_1 \leq t\} \cap \{\tau_2 > t\}} \frac{\partial_{u_1} P_t(\tau_1,T)}{\partial_{u_1} P_t(\tau_1,t)} \end{aligned}$$

The main observation here is that the price typically depends not only on the number of observed defaults so far, but also on the identities of the defaulted firms. This is reflected by the fact that the prices corresponding to the events $\{\tau_1 > t\} \cap \{\tau_2 \leq t\}$ and $\{\tau_1 \leq t\} \cap \{\tau_2 > t\}$, the second and third term above, are different in general. This should be contrasted to the case of ranked times (see e.g. [40], [56], [86]), where these two events cannot be distinguished. Compare Lemma 62, proved by El Karoui and al. in [40] with its slight generalization in Corollary 15. For pricing purposes, the ability to take the ordering of the default times into account does not provide any extra benefit, because only the unconditional expectation $P(\sigma_2 > T)$ is needed. However, in risk management, a typical task is to compute or simulate the distribution of the price, i.e. of $P(\sigma_2 > T \mid \mathcal{G}_t)$, at some future time $t > 0$. In this case the output is highly dependent on whether the random times were ranked or not, and this is where our results, for instance Corollary 15, adds to the existing literature.

We now proceed to make the comparison between the ranked and non-ranked case precise. To this end, let us introduce the filtration $\widetilde{\mathbb{G}} = (\widetilde{\mathcal{G}}_t)_{t\geq 0}$ given as the progressive expansion of $\mathbb{F}$ with the ranked times (σ_1, σ_2). It should be clear that $\widetilde{\mathbb{G}} \subset \mathbb{G}$. We are interested in $P(\sigma_2 > T \mid \widetilde{\mathcal{G}}_t)$ for $0 \leq t \leq T$, which we will compare with $P(\sigma_2 > T \mid \mathcal{G}_t)$ given in Corollary 15. The first step is to compute the $\mathcal{F}_t$-conditional distributions of (σ_1, σ_2) given those of (τ_1, τ_2).

Lemma 61 *Let $\widetilde{P}_t(u_1, u_2) = P(\sigma_1 > u_1, \sigma_2 > u_2 \mid \mathcal{F}_t)$. Then for $v_2 = u_1 \vee u_2$,*

$$\widetilde{P}_t(u_1, u_2) = P_t(u_1, v_2) + P_t(v_2, u_1) - P_t(v_2, v_2).$$

Proof. First note that for $u_2 \geq u_1$,

$$\begin{aligned} P(\sigma_1 > u_1, \sigma_2 \leq u_2 \mid \mathcal{F}_t) &= P(u_1 < \tau_1 \leq u_2, u_1 < \tau_2 \leq u_2 \mid \mathcal{F}_t) \\ &= P_t(u_1, u_1) - P_t(u_1, u_2) - P_t(u_2, u_1) + P_t(u_2, u_2). \end{aligned}$$

Now, since $\widetilde{P}_t(u_1, u_2) = \widetilde{P}_t(u_1, v_2) = P(\sigma_1 > u_1 \mid \mathcal{F}_t) - P(\sigma_1 > u_1, \sigma_2 \leq v_2 \mid \mathcal{F}_t)$ and $P(\sigma_1 > u_1 \mid \mathcal{F}_t) = P_t(u_1, u_1)$, the result follows. ■

Corollary 15 is a slight generalization of Lemma 62 below, and was obtained using the same methodology as in [40].

Lemma 62 (El karoui, Jeanblanc and Jiao) *Given the above notation and assumptions, we have*

$$\begin{aligned} P(\sigma_2 > T \mid \widetilde{\mathcal{G}}_t) = \mathbf{1}_{\{\sigma_1 > t\}} & \frac{P_t(T, t) + P_t(t, T) - P_t(T, T)}{P_t(t, t)} \\ & + \mathbf{1}_{\{\sigma_1 \leq t < \sigma_2\}} \frac{\partial_{u_1} P_t(\sigma_1, T) + \partial_{u_2} P_t(T, \sigma_1)}{\partial_{u_1} P_t(\sigma_1, t) + \partial_{u_2} P_t(t, \sigma_1)} \end{aligned}$$

Proof. An application of Corollary 15 with (τ_1, τ_2) replaced by (σ_1, σ_2), and $P_t(u_1, u_2)$ replaced by $\widetilde{P}_t(u_1, u_2)$ immediately yields

$$P(\sigma_2 > T \mid \widetilde{\mathcal{G}}_t) = \mathbf{1}_{\{\sigma_1 > t\}} \frac{\widetilde{P}_t(T, t) + \widetilde{P}_t(t, T) - \widetilde{P}_t(T, T)}{\widetilde{P}_t(t, t)} + \mathbf{1}_{\{\sigma_1 \leq t\} \cap \{\sigma_2 > t\}} \frac{\partial_{u_1} \widetilde{P}_t(\sigma_1, T)}{\partial_{u_1} \widetilde{P}_t(\sigma_1, t)}.$$

Using Lemma 61 to express $\widetilde{P}_t(u_1, u_2)$ in terms of $P_t(u_1, u_2)$ then yields the result. ■

Comparing $P(\sigma_2 > T \mid \widetilde{\mathcal{G}}_t)$ in Lemma 62 with the expression for $P(\sigma_2 > T \mid \mathcal{G}_t)$ in Corollary 15, we see that they coincide on the set $\{\sigma_1 > t\} = \{\tau_1 > t\} \cap \{\tau_2 > t\}$, as one would expect. However, as already mentioned, $\widetilde{\mathcal{G}}_t$ cannot distinguish between $\{\tau_1 >$

$t\} \cap \{\tau_2 \le t\}$ and $\{\tau_1 \le t\} \cap \{\tau_2 > t\}$, whose union is $\{\sigma_1 \le t < \sigma_2\}$. In fact, the value of $P(\sigma_2 > T \mid \widetilde{\mathcal{G}}_t)$ on this set is a conditional expectation of $P(\sigma_2 > T \mid \mathcal{G}_t)$. More specifically, one readily verifies that

$$\begin{aligned} P(\sigma_2 > T \mid \widetilde{\mathcal{G}}_t)\mathbf{1}_{\{\sigma_1 \le t < \sigma_2\}} = \alpha_t P(\sigma_2 > &T \mid \mathcal{G}_t)\mathbf{1}_{\{\tau_1 > t\}\cap\{\tau_2 \le t\}} \\ &+ (1-\alpha_t)P(\sigma_2 > T \mid \mathcal{G}_t)\mathbf{1}_{\{\tau_1 \le t\}\cap\{\tau_2 > t\}}, \end{aligned}$$

where

$$\alpha_t = \frac{\partial_{u_2} P_t(t,\sigma_1)}{\partial_{u_1} P_t(\sigma_1,t) + \partial_{u_2} P_t(t,\sigma_1)}.$$

In Section 2.3.4 we provide a simulation study to further illustrate the consequences of using $\widetilde{\mathbb{G}}$ as one's information set, rather than $\mathbb{G}$.

2.3.2 Extension to multiple non ranked random times

In this section we treat the general case of n default times, modeled as positive random variables $\tau_1, \dots, \tau_n$, not necessarily ordered extending thereby a bit the results in [40] and [41]. We derive intensities and pricing formulas under the assumption that joint $\mathcal{F}_t$-conditional densities exist, using the methodology developed in [40]. The market information set $\mathbb{G}$ is the progressive expansion of $\mathbb{F}$ with $\tau_1, \dots, \tau_n$.

The treatment requires some notation, which we now describe. For a vector $\boldsymbol{x} \in \mathbb{R}^n$ and an index set $I \subset \{1, \dots, n\}$, let $\boldsymbol{x}_I$ be the subvector whose entries are the components of $\boldsymbol{x}$ with indices in I. In particular, we define $\boldsymbol{\tau} = (\tau_1, \dots, \tau_n)$ so that $\boldsymbol{\tau}_I = (\tau_i)_{i\in I}$. Inequalities like $\boldsymbol{x}_I > t$ for a vector $\boldsymbol{x}$, an index set $I \subset \{1, \dots, n\}$, and a scalar t should be interpreted as $x_i > t$ for all $i \in I$. In particular,

$$\mathbf{1}_{\{\boldsymbol{\tau}_I > t\}} = \mathbf{1}_{\{\tau_i > t,\ \forall i \in I\}}.$$

As a convention, when $I = \emptyset$, $\mathbf{1}_{\{\boldsymbol{\tau}_I > t\}} = 1$ and $\mathbf{1}_{\{\boldsymbol{\tau}_I \le t\}} = 1$. We take $t \wedge \boldsymbol{x}_I$ to denote the vector $(t \wedge x_i)_{i\in I}$, and let $\boldsymbol{e} = (1, \dots, 1)$ be the vector consisting of only ones. Its dimension should be clear from the context.

Several different filtrations will appear. For an index set $I \subset \{1, \dots, n\}$, let $\mathbb{F}^I$ denote the progressive expansion of $\mathbb{F}$ with all the times $\boldsymbol{\tau}_I$. The market filtration $\mathbb{G}$, the full filtration containing all the times $\tau_1, \dots, \tau_n$, is then given by $\mathbb{G} = \mathbb{F}^{\{1,\dots,n\}}$. We also write $\mathbb{G}^{-k} = \mathbb{F}^{\{1,\dots,n\}\setminus\{k\}}$, a filtration which will play an important role.

Throughout this section we assume that $\mathcal{F}_t$-conditional densities exist. More precisely, we assume that there exist processes $(p_t(\boldsymbol{u}))_{t\ge 0}$ indexed by $\boldsymbol{u} \in \mathbb{R}^n_+$ such that

$$P(\boldsymbol{\tau} \in d\boldsymbol{u} \mid \mathcal{F}_t) = p_t(\boldsymbol{u})d\boldsymbol{u}.$$

It is not essential that these densities exist with respect to Lebesgue measure; any other deterministic and diffuse measure would suffice. We use Lebesgue measure primarily in order to lighten the already quite complicated notation. The key result is the following theorem. In its statement, the measure $\mu(dx)$ is some given deterministic measure.

Theorem 44 *Let X be an integrable random variable, and let I and J be disjoint subsets of $\{1,\ldots,n\}$. Suppose that for every $t \geq 0$, $(X, \boldsymbol{\tau}_I, \boldsymbol{\tau}_J)$ has a joint $\mathcal{F}_t$-conditional density $f_t(x, \boldsymbol{u}_I, \boldsymbol{u}_J)$ with respect to $\mu(dx) \times d\boldsymbol{u}_I \times d\boldsymbol{u}_J$. Then, on $\{\boldsymbol{\tau}_I > t\} \cap \{\boldsymbol{\tau}_J \leq t\}$,*

$$E(X \mid \mathcal{F}_t^{I\cup J}) = \int_{\mathbb{R}} x g_t(x \mid I, J)\mu(dx),$$

where

$$g_t(x \mid I, J) = \frac{\int_t^\infty \ldots \int_t^\infty f_t(x, \boldsymbol{u}_I, \boldsymbol{\tau}_J) d\boldsymbol{u}_I}{\int_{\mathbb{R}} \int_t^\infty \ldots \int_t^\infty f_t(x, \boldsymbol{u}_I, \boldsymbol{\tau}_J) d\boldsymbol{u}_I \mu(dx)}.$$

Despite its cumbersome appearance, the function $g_t(x \mid I, J)$ appearing in Theorem 44 has a very natural interpretation. It is the conditional density of X with respect to $\mu(dx)$, given $\mathcal{F}_t$, $\{\boldsymbol{\tau}_I > t\}$, and $\boldsymbol{\tau}_J$. Formally, we may write

$$g_t(x \mid I, J)\mu(dx) = P(X \in dx \mid \mathcal{F}_t,\ \boldsymbol{\tau}_I > t,\ \boldsymbol{\tau}_J).$$

Notice also that the sets I or J (or both) could be empty. The statement is still valid, as long as the correct notational conventions are adhered to. Namely, $\mathbf{1}_{\{\boldsymbol{\tau}_I > t\}}$ and $\mathbf{1}_{\{\boldsymbol{\tau}_J \leq t\}}$ are equal to one if $I = \emptyset$ and $J = \emptyset$, respectively. Also, integrals with respect to $d\boldsymbol{u}_I$ (respectively $d\boldsymbol{u}_J$) are ignored in this case. We prove now Theorem 44. In order to prove this theorem, it is convenient to start with the following weaker result.

Lemma 63 *Let X be an integrable random variable, and let I and J be disjoint subsets of $\{1,\ldots,n\}$. Suppose that for every $t \geq 0$, $(X, \boldsymbol{\tau}_I, \boldsymbol{\tau}_J)$ has a joint $\mathcal{F}_t$-conditional density $f_t(x, \boldsymbol{u}_I, \boldsymbol{u}_J)$ with respect to $\mu(dx) \times d\boldsymbol{u}_I \times d\boldsymbol{u}_J$. Then, on $\{\boldsymbol{\tau}_I > t\} \cap \{\boldsymbol{\tau}_J \leq t\}$,*

$$E(X \mid \mathcal{F}_t^{I\cup J}) = \frac{1}{P(\boldsymbol{\tau}_I > t \mid \mathcal{F}_t^J)} \frac{\int_t^\infty \ldots \int_t^\infty \int_{\mathbb{R}} x f_t(x, \boldsymbol{u}_I, \boldsymbol{\tau}_J)\mu(dx) d\boldsymbol{u}_I}{\int_0^\infty \ldots \int_0^\infty \int_{\mathbb{R}} f_t(x, \boldsymbol{u}_I, \boldsymbol{\tau}_J)\mu(dx) d\boldsymbol{u}_I}.$$

Proof. Let $U = Y_t h_I(t \wedge \boldsymbol{\tau}_I) h_J(t \wedge \boldsymbol{\tau}_J)$, where Y_t is a bounded $\mathcal{F}_t$ measurable random variable, and $h_I : \mathbb{R}^{|I|} \to \mathbb{R}$ and $h_J : \mathbb{R}^{|J|} \to \mathbb{R}$ are bounded and measurable. We have

$$E(U \mathbf{1}_{\{\boldsymbol{\tau}_I > t\}} \mathbf{1}_{\{\boldsymbol{\tau}_J \leq t\}} X) = E(Y_t h_I(t\boldsymbol{e}) E(X \mathbf{1}_{\{\boldsymbol{\tau}_J \leq t\}} h_J(\boldsymbol{\tau}_J) \mathbf{1}_{\{\boldsymbol{\tau}_I > t\}} \mid \mathcal{F}_t)),$$

(where $t\boldsymbol{e} = (t, \cdots, t)$) and

$$\begin{aligned}
&E(X \mathbf{1}_{\{\boldsymbol{\tau}_J \leq t\}} h_J(\boldsymbol{\tau}_J) \mathbf{1}_{\{\boldsymbol{\tau}_I > t\}} \mid \mathcal{F}_t) \\
&\quad = \int_0^\infty \cdots \int_0^\infty \int_{\mathbb{R}} x \mathbf{1}_{\{\boldsymbol{u}_J \leq t\}} h_J(\boldsymbol{u}_J) \mathbf{1}_{\{\boldsymbol{u}_I > t\}} f_t(x, \boldsymbol{u}_I, \boldsymbol{u}_J)\mu(dx) d\boldsymbol{u}_I d\boldsymbol{u}_J \\
&\quad = \int_0^\infty \cdots \int_0^\infty \mathbf{1}_{\{\boldsymbol{u}_J \leq t\}} h_J(\boldsymbol{u}_J) \left(\int_0^\infty \cdots \int_0^\infty \int_{\mathbb{R}} x \mathbf{1}_{\{\boldsymbol{u}_I > t\}} f_t(x, \boldsymbol{u}_I, \boldsymbol{u}_J)\mu(dx) d\boldsymbol{u}_I \right) d\boldsymbol{u}_J \\
&\quad = \int_0^\infty \cdots \int_0^\infty \mathbf{1}_{\{\boldsymbol{u}_J \leq t\}} h_J(\boldsymbol{u}_J) G(\boldsymbol{u}_J) \int_0^\infty \cdots \int_0^\infty \int_{\mathbb{R}} f_t(x, \boldsymbol{u}_I, \boldsymbol{u}_J)\mu(dx) d\boldsymbol{u}_I d\boldsymbol{u}_J \\
&\quad = E(\mathbf{1}_{\{\boldsymbol{\tau}_J \leq t\}} h(\boldsymbol{\tau}_J) G(\boldsymbol{\tau}_J) \mid \mathcal{F}_t),
\end{aligned}$$

where, for this proof only, we use the short-hand notation

$$G(\boldsymbol{u}_J) = \frac{\int_t^\infty \cdots \int_t^\infty \int_{\mathbb{R}} x f_t(x, \boldsymbol{u}_I, \boldsymbol{u}_J)\mu(dx)d\boldsymbol{u}_I}{\int_0^\infty \cdots \int_0^\infty \int_{\mathbb{R}} f_t(x, \boldsymbol{u}_I, \boldsymbol{u}_J)\mu(dx)d\boldsymbol{u}_I}.$$

Hence $E(U\mathbf{1}_{\{\boldsymbol{\tau}_I>t\}}\mathbf{1}_{\{\boldsymbol{\tau}_J\le t\}}X) = E(Y_t h_I(t\boldsymbol{e})h_J(\boldsymbol{\tau}_J)\mathbf{1}_{\{\boldsymbol{\tau}_J\le t\}}G(\boldsymbol{\tau}_J))$, and continuing the calculations we obtain

$$\begin{aligned}
&E(Y_t h_I(t\boldsymbol{e})h_J(\boldsymbol{\tau}_J)\mathbf{1}_{\{\boldsymbol{\tau}_J\le t\}}G(\boldsymbol{\tau}_J)) \\
&\qquad = E\left\{Y_t h_I(t\boldsymbol{e})\frac{E(\mathbf{1}_{\{\boldsymbol{\tau}_I>t\}} \mid \mathcal{F}_t^J)}{P(\boldsymbol{\tau}_I > t \mid \mathcal{F}_t^J)}h_J(\boldsymbol{\tau}_J)\mathbf{1}_{\{\boldsymbol{\tau}_J\le t\}}G(\boldsymbol{\tau}_J)\right\} \\
&\qquad = E\left\{Y_t h_I(t\boldsymbol{e})h_J(\boldsymbol{\tau}_J)\mathbf{1}_{\{\boldsymbol{\tau}_I>t\}}\mathbf{1}_{\{\boldsymbol{\tau}_J\le t\}}\frac{1}{P(\boldsymbol{\tau}_I > t \mid \mathcal{F}_t^J)}G(\boldsymbol{\tau}_J)\right\} \\
&\qquad = E\left\{U\mathbf{1}_{\{\boldsymbol{\tau}_I>t\}}\mathbf{1}_{\{\boldsymbol{\tau}_J\le t\}}\frac{1}{P(\boldsymbol{\tau}_I > t \mid \mathcal{F}_t^J)}G(\boldsymbol{\tau}_J)\right\}.
\end{aligned}$$

The Monotone Class Theorem and a standard right-continuity argument now yields the result. ■

The fact that the statement and proof of Lemma 63 remain valid even if I and/or J is empty allows us to improve its conclusion without any additional assumptions. The result is Theorem 44, and we now give the proof.

Proof of Theorem 44. Applying Lemma 63 with $I = \emptyset$ and an integrable random variable $\tilde{X}$ yields, with $\tilde{f}_t(x, \boldsymbol{u}_J)$ being the conditional joint density of $(\tilde{X}, \boldsymbol{\tau}_J)$ with respect to $\tilde{\mu}(dx) \times d\boldsymbol{u}_J$,

$$\mathbf{1}_{\{\boldsymbol{\tau}_J\le t\}}E(\tilde{X} \mid \mathcal{F}_t^J) = \mathbf{1}_{\{\boldsymbol{\tau}_J\le t\}}\frac{\int_{\mathbb{R}} x\tilde{f}_t(x, \boldsymbol{\tau}_J)\tilde{\mu}(dx)}{\int_{\mathbb{R}} \tilde{f}_t(x, \boldsymbol{\tau}_J)\tilde{\mu}(dx)}.$$

In this expression we would like to take $\tilde{X} = \mathbf{1}_{\{\boldsymbol{\tau}_I>t\}}$, with I as in the statement of the theorem. Taking $\tilde{\mu}(dx) = \delta_{\{0,1\}}(dx)$, the conditional joint density $\tilde{f}_t(x, \boldsymbol{u}_J)$ can be computed in terms of $f_t(x, \boldsymbol{u}_I, \boldsymbol{u}_J)$. We get

$$\begin{aligned}
\tilde{f}_t(0, \boldsymbol{u}_J) &= \int_0^t \cdots \int_0^t \int_{\mathbb{R}} f_t(x, \boldsymbol{u}_I, \boldsymbol{u}_J)\mu(dx)d\boldsymbol{u}_I, \\
\tilde{f}_t(1, \boldsymbol{u}_J) &= \int_t^\infty \cdots \int_t^\infty \int_{\mathbb{R}} f_t(x, \boldsymbol{u}_I, \boldsymbol{u}_J)\mu(dx)d\boldsymbol{u}_I.
\end{aligned}$$

A straightforward calculation then gives

$$\mathbf{1}_{\{\boldsymbol{\tau}_J\le t\}}E(\mathbf{1}_{\{\boldsymbol{\tau}_I>t\}} \mid \mathcal{F}_t^J) = \mathbf{1}_{\{\boldsymbol{\tau}_J\le t\}}\frac{\int_t^\infty \cdots \int_t^\infty \int_{\mathbb{R}} f_t(x, \boldsymbol{u}_I, \boldsymbol{\tau}_J)\mu(dx)d\boldsymbol{u}_I}{\int_0^\infty \cdots \int_0^\infty \int_{\mathbb{R}} f_t(x, \boldsymbol{u}_I, \boldsymbol{\tau}_J)\mu(dx)d\boldsymbol{u}_I}.$$

It now suffices to apply Lemma 63 with I, J and X as in the theorem, and substitute for $\mathbf{1}_{\{\boldsymbol{\tau}_J\le t\}}E(\mathbf{1}_{\{\boldsymbol{\tau}_I>t\}} \mid \mathcal{F}_t^J)$, to get

$$\mathbf{1}_{\{\boldsymbol{\tau}_I>t\}}\mathbf{1}_{\{\boldsymbol{\tau}_J\le t\}}E(X \mid \mathcal{F}_t^{I\cup J}) = \mathbf{1}_{\{\boldsymbol{\tau}_I>t\}}\mathbf{1}_{\{\boldsymbol{\tau}_J\le t\}}\frac{\int_t^\infty \cdots \int_t^\infty \int_{\mathbb{R}} x f_t(x, \boldsymbol{u}_I, \boldsymbol{\tau}_J)\mu(dx)d\boldsymbol{u}_I}{\int_t^\infty \cdots \int_t^\infty \int_{\mathbb{R}} f_t(x, \boldsymbol{u}_I, \boldsymbol{\tau}_J)\mu(dx)d\boldsymbol{u}_I},$$

which is the desired result. ■

Using Theorem 44 we can compute, for every fixed k, the $\mathbb{G}^{-k}$ supermartingale

$$Z_t^k = P(\tau_k > t \mid \mathcal{G}_t^{-k}),$$

together with its Doob-Meyer decomposition. (Recall that $\mathbb{G}^{-k} = \mathbb{F}^{\{1,\ldots,n\}\backslash\{k\}}$ is the progressive expansion of $\mathbb{F}$ with all the times except τ_k.) This will then allow us to find the compensator and default intensity of τ_k. First we introduce the following notation, analogous to the definition of $g_t(x \mid I, J)$ appearing in Theorem 44. Define

$$\beta_t(u \mid I, J) = \frac{\int_t^\infty \cdots \int_t^\infty p_t(\boldsymbol{u}_I, \boldsymbol{\tau}_J, u) d\boldsymbol{u}_I}{\int_0^\infty \int_t^\infty \cdots \int_t^\infty p_t(\boldsymbol{u}_I, \boldsymbol{\tau}_J, u_k) d\boldsymbol{u}_I du_k}, \tag{2.5}$$

where of course $p_t(\boldsymbol{u})$ is the conditional joint density of $\tau_1, \ldots, \tau_n$. In particular, when $J = \emptyset$ this reduces to

$$\beta_t(u \mid I, \emptyset) = \frac{\int_t^\infty \cdots \int_t^\infty p_t(\boldsymbol{u}_I, u) d\boldsymbol{u}_I}{\int_0^\infty \int_t^\infty \cdots \int_t^\infty p_t(\boldsymbol{u}_I, u_k) d\boldsymbol{u}_I du_k}.$$

Note that $\beta_t(u \mid I, J)$ is the conditional density of τ_k given $\mathcal{F}_t$, $\{\boldsymbol{\tau}_I > t\}$ and $\boldsymbol{\tau}_J$. We write this formally as

$$\beta_t(u \mid I, J) du = P(\tau_k \in du \mid \mathcal{F}_t,\ \boldsymbol{\tau}_I > t,\ \boldsymbol{\tau}_J).$$

The following two lemmas are the key ingredients for computing the $\mathbb{G}$ intensity of τ_k.

Lemma 64 *Fix $k \in \{1, \ldots, n\}$ and let $\beta_t(u \mid I, J)$ be defined as in (2.5). Then*

$$P(\tau_k > T \mid \mathcal{G}_t^{-k}) = \sum_{I,J} \mathbf{1}_{\{\boldsymbol{\tau}_I > t\} \cap \{\boldsymbol{\tau}_J \le t\}} \int_T^\infty \beta_t(u \mid I, J) du,$$

where the sum is taken over all subsets $I, J \subset \{1, \ldots, n\}$ such that $I \cup J = \{1, \ldots, n\} \backslash \{k\}$ and $I \cap J = \emptyset$.

Proof. Let $T > 0$. With $X = \mathbf{1}_{\{\tau_k > T\}}$, $\mu(dx) = \delta_{\{0,1\}}(dx)$ and

$$f_t(1, \boldsymbol{u}_I, \boldsymbol{u}_J) = \int_T^\infty \beta_t(u \mid I, J; \boldsymbol{u}_J) du,$$

where $\beta_t(u \mid I, J; \boldsymbol{u}_J)$ is given by the right side of (2.5) but with $\boldsymbol{\tau}_J$ replaced by $\boldsymbol{u}_J$, an application of Theorem 44 yields

$$\begin{aligned} P(\tau_k > T \mid \mathcal{G}_t^{-k}) &= \sum_{I,J} \mathbf{1}_{\{\boldsymbol{\tau}_I > t\}} \mathbf{1}_{\{\boldsymbol{\tau}_J \le t\}} P(\tau_k > T \mid \mathcal{G}_t^{-k}) \\ &= \sum_{I,J} \mathbf{1}_{\{\boldsymbol{\tau}_I > t\}} \mathbf{1}_{\{\boldsymbol{\tau}_J \le t\}} \int_T^\infty \beta_t(u \mid I, J; \boldsymbol{\tau}_J) du, \end{aligned}$$

where the sums are taken over all $I, J \subset \{1, \ldots, n\}$ with $I \cup J = \{1, \ldots, n\} \backslash \{k\}$ and $I \cap J = \emptyset$. Since $\beta_t(u \mid I, J; \boldsymbol{\tau}_J) = \beta_t(u \mid I, J)$, the result follows. ■

Lemma 65 *Fix $k \in \{1, \ldots, n\}$ and let $Z^k = M^k - A^k$ be the Doob-Meyer decomposition in $\mathbb{G}^{-k}$ of Z^k. On $\{\boldsymbol{\tau}_I > t\} \cap \{\boldsymbol{\tau}_J \leq t\}$ we then have*

$$M_t^k = \int_0^\infty \beta_{u \wedge t}(u \mid I, J) du \qquad \text{and} \qquad A_t^k = \int_0^t \beta_u(u \mid I, J) du,$$

where $I, J \subset \{1, \ldots, n\}$ satisfy $I \cup J = \{1, \ldots, n\} \backslash \{k\}$ and $I \cap J = \emptyset$.

Proof. Lemma 64 implies that $P(\tau_k > T \mid \mathcal{G}_t^{-k}) = \int_T^\infty \alpha_t(u) du$, where

$$\alpha_t(u) = \sum_{I,J} \mathbf{1}_{\{\boldsymbol{\tau}_I > t\}} \mathbf{1}_{\{\boldsymbol{\tau}_J \leq t\}} \beta_t(u \mid I, J).$$

Since $P(\tau_k > T \mid \mathcal{G}_t^{-k})$ is a martingale, so is $\alpha_t(u)$ for every u. We write

$$Z_t^k = \int_t^\infty \alpha_t(u) du = \int_0^\infty \alpha_{u \wedge t}(u) du - \int_0^t \alpha_u(u) du,$$

and notice that the first term on the right is a martingale and the second is of finite variation and continuous, hence predictable. This gives the desired result. The above computation mimics Proposition 2.1 in [86]. ■

The proof of Theorem 45 below relies on the Jeulin-Yor Theorem, see Theorem 7. We may finally give the intensities explicitly in terms of the $\mathcal{F}_t$-conditional densities $p_t(\boldsymbol{u})$, or, more precisely, in terms of the $\beta_t(u \mid I, J)$ defined in (2.5).

Theorem 45 (El Karoui, Jeanblanc and Jiao) *For each fixed k, the process*

$$\mathbf{1}_{\{\tau_k \leq t\}} - \int_0^{t \wedge \tau_k} \lambda_s^k ds$$

is a $\mathbb{G}$ martingale, where

$$\lambda_t^k = \frac{\beta_t(t \mid I, J)}{\int_t^\infty \beta_t(u \mid I, J) du} \qquad \text{on} \quad \{\boldsymbol{\tau}_I > t\} \cap \{\boldsymbol{\tau}_J \leq t\}.$$

Here $\beta_t(u \mid I, J)$ is given by (2.5), and $I, J \subset \{1, \ldots, n\}$ satisfy $I \cup J = \{1, \ldots, n\} \backslash \{k\}$ and $I \cap J = \emptyset$.

This theorem is actually a slight generalization of the one proved by El Karoui, Jeanblanc and Jiao in the case of ranked times (see [40]). We used their exact methodology and our result is similar to theirs up to the cumbersome combinatorial effect. However the formulas of Theorem 45 are needed both from a risk management perspective as this will be illustrated in section 2.3.4 and to study credit contagion as illustrated in sections 2.3.5 and 2.3.6. Let $(\sigma_i)_{1 \leq i \leq n}$ be the ordered statistics of the default times $(\tau_i)_{1 \leq i \leq n}$ and let $\tilde{\mathbb{G}}$ be the progressive expansion of $\mathbb{F}$ with $(\sigma_i)_{1 \leq i \leq n}$. Finally, let $\tilde{p}$ be the $\mathcal{F}_t$-conditional density of $(\sigma_1, \ldots, \sigma_n)$ and $\tilde{\beta}$ the corresponding quantities to those in equation (2.5) by replacing p with $\tilde{p}$. Since the times are ordered now, $\tilde{\beta}_t(u \mid I, J)$ does not vanish if and only if the times indexed by J are smaller than u which is itself smaller than the times indexed by I. This leaves only one term and gives the following.

Theorem 46 (El Karoui, Jeanblanc and Jiao) *For each fixed k, the process*

$$\mathbf{1}_{\{\sigma_k \le t\}} - \int_0^{t\wedge\sigma_k} \tilde{\lambda}_s^k ds$$

is a $\tilde{\mathbb{G}}$ martingale, where

$$\tilde{\lambda}_t^k = 1_{\{\sigma_{k-1}<t\}} \frac{\tilde{\beta}_t(t \mid \{k+1,\dots,n\},\{1,\dots,k-1\})}{\int_t^\infty \tilde{\beta}_t(u \mid \{k+1,\dots,n\},\{1,\dots,k-1\})du}$$

As expected, the conclusion of Theorem 45 may formally be written

$$\lambda_t^k dt = P(\tau_k \in dt \mid \mathcal{F}_t,\ \boldsymbol{\tau}_I > t,\ \tau_k > t,\ \boldsymbol{\tau}_J).$$

Proof. [Of Theorem 45] Lemma 64 gives us the form of Z^k, and Lemma 65 provides its Doob-Meyer decomposition $Z^k = M^k - A^k$ in $\mathbb{G}^{-k}$. An application of the Jeulin-Yor Theorem (Theorem 7) with the filtrations $\mathbb{G}^{-k} \subset \mathbb{G}$ now gives the result. Notice that since A^k from Lemma 65 is continuous, there is no need to take left limits of Z^k when applying the Jeulin-Yor Theorem. ■

By substituting for $\beta_t(u \mid I, J)$, we may also write λ_t^k directly in terms of $p_t(\boldsymbol{u})$:

$$\lambda_t^k = \frac{\int_t^\infty \cdots \int_t^\infty p_t(\boldsymbol{u}_I, \boldsymbol{\tau}_J, t) d\boldsymbol{u}_I}{\int_t^\infty \cdots \int_t^\infty p_t(\boldsymbol{u}_I, \boldsymbol{\tau}_J, u) d\boldsymbol{u}_I du} \qquad \text{on} \quad \{\boldsymbol{\tau}_I > t\} \cap \{\boldsymbol{\tau}_J \le t\}. \tag{2.6}$$

This form makes it clear that the intensity of τ_k will jump upon the arrival of some other default τ_i, unless the conditional densities have some special structure. As described in Section 2.3.1, the interpretation of this effect is that the default of firm i gives indirect information about the state of firm k, thereby causing its intensity to jump. As we will see in section 2.3.5, one case where this type of contagion does not occur is when conditional independence is present, i.e. when the conditional densities factor out.

We provide now pricing formulas with respect to the market filtration. The risk-neutral pricing technology reduces the problem of pricing payoffs obtained at time T to the problem of computing the expectation of the (discounted) payoff under a risk-neutral measure. The price at time $t < T$ is given by the conditional expectation given the market information $\mathcal{G}_t$. We show how this can be done in the present framework, first for general claims with a fixed maturity, and then for the special case of k^{th}-to-default swaps. The next result is standard for the case $n = 1$, and nothing essential changes in the proof when $n > 1$. Nonetheless, we give the details for completeness.

Lemma 66 *Let X be an integrable random variable. Fix $k \in \{1,\dots,n\}$ and define $Z_t^k = P(\tau_k > t \mid \mathcal{G}_t^{-k})$. Then for $T \ge t$,*

$$E(X\mathbf{1}_{\{\tau_k > T\}} \mid \mathcal{G}_t) = \mathbf{1}_{\{\tau_k > t\}} \frac{1}{Z_t^k} E(X\mathbf{1}_{\{\tau_k > T\}} \mid \mathcal{G}_t^{-k}).$$

Moreover, if X is $\mathcal{G}_T^{-k}$ measurable, then

$$E(X\mathbf{1}_{\{\tau_k>T\}} \mid \mathcal{G}_t) = \mathbf{1}_{\{\tau_k>t\}} \frac{1}{Z_t^k} E(XZ_T^k \mid \mathcal{G}_t^{-k}).$$

Proof. Let U be $\mathcal{G}_t$ measurable of the form $U = Y_t h_{-k}(t \wedge \boldsymbol{\tau}_{-k}) h_k(t \wedge \tau_k)$, where Y_t is $\mathcal{F}_t$ measurable and $h_{-k} : \mathbb{R}^{n-1} \to \mathbb{R}$ and $h_k : \mathbb{R} \to \mathbb{R}$ are bounded Borel functions. Then

$$\begin{aligned}
E(X\mathbf{1}_{\{\tau_k>T\}}U) &= E(Y_t h_k(t) h_{-k}(t \wedge \boldsymbol{\tau}_{-k}) X\mathbf{1}_{\{\tau_k>T\}}) \\
&= E(Y_t h_k(t) h_{-k}(t \wedge \boldsymbol{\tau}_{-k}) E(X\mathbf{1}_{\{\tau_k>T\}} \mid \mathcal{G}_t^{-k})) \\
&= E\left\{ Y_t h_k(t) h_{-k}(t \wedge \boldsymbol{\tau}_{-k}) \frac{E(\mathbf{1}_{\{\tau_k>t\}} \mid \mathcal{G}_t^{-k})}{Z_t^k} E(X\mathbf{1}_{\{\tau_k>T\}} \mid \mathcal{G}_t^{-k}) \right\} \\
&= E\left\{ E\left(U \frac{\mathbf{1}_{\{\tau_k>t\}} E(X\mathbf{1}_{\{\tau_k>T\}} \mid \mathcal{G}_t^{-k})}{Z_t^k} \mid \mathcal{G}_t^{-k} \right) \right\} \\
&= E\left\{ U \frac{\mathbf{1}_{\{\tau_k>t\}} E(X\mathbf{1}_{\{\tau_k>T\}} \mid \mathcal{G}_t^{-k})}{Z_t^k} \right\}
\end{aligned}$$

An application of the Monotone Class Theorem together with a standard right-continuity argument proves the first claim of the lemma. If X is $\mathcal{G}_T^{-k}$ measurable, simply condition on $\mathcal{G}_T^{-k}$ to get $E(X\mathbf{1}_{\{\tau_k>T\}} \mid \mathcal{G}_t^{-k}) = E(XZ_T^k \mid \mathcal{G}_t^{-k})$. ■

Under suitable assumptions about existence of joint $\mathcal{F}_t$-conditional densities, Theorem 44 allows to compute the expectation $E(X\mathbf{1}_{\{\tau_k>T\}} \mid \mathcal{G}_t^{-k})$ that appears in Lemma 66. Moreover, Z_t^k is given by Lemma 64.

As a more specific example, we give a formula for the price of a security that pays one unit of currency if at most $k-1$ defaults happen before some fixed horizon T. This quantity appears if one wants to price k^{th}-to-default swaps. We assume that P is a valid pricing measure, and that interest rates are zero (this is without loss of generality using a change of numeraire.) Let

$$\sigma_1 \leq \sigma_2 \leq \cdots \leq \sigma_n$$

be the order statistics of the default times $\tau_1, \ldots, \tau_n$. That is,

$$\sigma_1 = \min\{\tau_1, \ldots, \tau_n\}, \qquad \sigma_2 = \min\{\tau_1, \ldots, \tau_n\} \backslash \{\sigma_1\}, \qquad \text{etc.}$$

The time t price of the security described above is then given by the following.

Lemma 67 *For any $t \leq T$,*

$$\begin{aligned}
P(\sigma_k > T \mid \mathcal{G}_t) = \sum_{|I| \geq k} \sum_{K \supset I} & \mathbf{1}_{\{\boldsymbol{\tau}_K > t\} \cap \{\boldsymbol{\tau}_{K^c} \leq t\}} \\
& \times \frac{\int_T^\infty \cdots \int_T^\infty \int_t^T \cdots \int_t^T p_t(\boldsymbol{u}_I, \boldsymbol{u}_{I^c \backslash K^c}, \boldsymbol{\tau}_{K^c}) d\boldsymbol{u}_{I^c \backslash K^c} d\boldsymbol{u}_I}{\int_t^\infty \cdots \int_t^\infty p_t(\boldsymbol{u}_K, \boldsymbol{\tau}_{K^c}) d\boldsymbol{u}_K}
\end{aligned}$$

Proof. Note that $\mathbf{1}_{\{\sigma_k>T\}} = \sum_{|I|\geq k} \mathbf{1}_{\{\boldsymbol{\tau}_I>T\}}\mathbf{1}_{\{\boldsymbol{\tau}_J\leq T\}}$, where here and in the rest of this proof, $J = I^c$. Now,

$$\begin{aligned}P(\boldsymbol{\tau}_I > T,\ \boldsymbol{\tau}_J \leq T \mid \mathcal{G}_t) &= \sum_{K\supset I} \mathbf{1}_{\{\boldsymbol{\tau}_K>t\}\cap\{\boldsymbol{\tau}_{K^c}\leq t\}} P(\boldsymbol{\tau}_I > T,\ \boldsymbol{\tau}_J \leq T \mid \mathcal{G}_t)\\ &= \sum_{K\supset I} \mathbf{1}_{\{\boldsymbol{\tau}_K>t\}\cap\{\boldsymbol{\tau}_{K^c}\leq t\}} P(\boldsymbol{\tau}_I > T,\ t \leq \boldsymbol{\tau}_{J\setminus K^c} \leq T \mid \mathcal{G}_t).\end{aligned}$$

We claim that on $\{\boldsymbol{\tau}_K > t\} \cap \{\boldsymbol{\tau}_{K^c} \leq t\}$,

$$P(\boldsymbol{\tau}_I > T,\ t \leq \boldsymbol{\tau}_{J\setminus K^c} \leq T \mid \mathcal{G}_t) = \frac{\int_T^\infty \cdots \int_T^\infty \int_t^T \cdots \int_t^T p_t(\boldsymbol{u}_I, \boldsymbol{u}_{J\setminus K^c}, \boldsymbol{\tau}_{K^c}) d\boldsymbol{u}_{J\setminus K^c} d\boldsymbol{u}_I}{\int_t^\infty \cdots \int_t^\infty p_t(\boldsymbol{u}_I, \boldsymbol{u}_{J\setminus K^c}, \boldsymbol{\tau}_{K^c}) d\boldsymbol{u}_{J\setminus K^c} d\boldsymbol{u}_I}.$$

The cumbersome verification of this fact involves the same Monotone Class and right-continuity arguments used before. The result now follows by assembling the pieces. ■

The expression for $P(\sigma_k > T \mid \mathcal{G}_t)$ is quite complicated, as it involves a potentially very large number of terms, where each term consists of a high-dimensional integral. Such quantities are not tractable in general.

2.3.3 Modeling of conditional densities

Until now we have taken the vector of default times *and* the $\mathcal{F}_t$-conditional densities as given. However, from a modeling perspective it is more convenient to start by specifying the conditional densities, and then construct the random times to be consistent with this specification. We now show how this can be done. We will then provide ways to model the conditional densities. For the construction and simulation of the random time, we will in fact treat a more general case, where τ is a random variable with values in some measurable space $(\mathbb{E}, \mathcal{E})$. Of course, when τ is a vector of random times, $\mathbb{E} = \mathbb{R}_+^n$.

Let us assume that we are given a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ and a state space $\mathbb{E}$ with σ-field $\mathcal{E}$, together with processes $(\omega, t) \mapsto P_t(E; \omega)$ indexed by the sets $E \in \mathcal{E}$. We assume that for all t and P-a.e. ω, $P_t(\cdot) = P_t(\cdot; \omega)$ is a probability measure on $\mathcal{E}$. We also assume that the process $(P_t(E))_{t\geq 0}$ is a uniformly integrable $\mathbb{F}$ martingale for every $E \in \mathcal{E}$, and we define the random set function $\mu(\cdot) = \mu(\cdot; \omega)$ by

$$\mu(E; \omega) = P_\infty(E; \omega) = \lim_{t\to\infty} P_t(E; \omega).$$

This is well-defined outside a P-nullset. Finally, we assume that $\mu(\cdot; \omega)$ is indeed a probability measure a.s. This assumption is needed since countable additivity can fail if the limits $\lim_{t\to\infty} P_t(E; \omega)$ are not uniform across $E \in \mathcal{E}$. One situation where countable additivity automatically holds is if all the $P_t(E)$ are constant after some fixed time T. The processes $(P_t(E))_{t\geq 0}$ should be thought of as candidates for a system of $\mathcal{F}_t$-conditional probabilities associated with some $\mathbb{E}$-valued random variable τ. The goal is to construct this random variable, possibly on some extension $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{P})$ of the original space. This can be done explicitly. Define

$$\tilde{\Omega} = \Omega \times \mathbb{E} \qquad \tilde{\mathcal{F}} = \mathcal{F} \otimes \mathcal{E} \qquad \tilde{\mathcal{F}}_t = \mathcal{F}_t \otimes \{\emptyset, \mathbb{E}\},$$

and let $\tilde{P}$ be the unique measure on $\tilde{\mathcal{F}}$ that satisfies

$$\tilde{P}(A \times E) = \int_A \mu(E;\omega) dP(\omega)$$

for every $A \in \mathcal{F}$ and $E \in \mathcal{E}$. In particular, this implies that the projection of $\tilde{P}$ onto Ω is P. That is,

$$\tilde{P}(\cdot \times \mathbb{E}) = P(\cdot). \tag{2.7}$$

Mappings $\omega \mapsto X(\omega)$ on Ω are naturally identified with mappings on $\tilde{\Omega}$ by setting $X(\tilde{\omega}) = X(\omega)$ for $\tilde{\omega} = (\omega, e)$. Measurability properties carry over: if $X(\omega)$ is $\mathcal{F}$ measurable, then $X(\tilde{\omega})$ is $\tilde{\mathcal{F}}$ measurable; if $X(\omega)$ is an $\mathbb{F}$ adapted process, then $X(\tilde{\omega})$ is $\tilde{\mathbb{F}}$ adapted, etc. Moreover, it follows from (2.7) that the law under P of X (considered as a mapping on Ω) is the same as its law under $\tilde{P}$ (when considered as a mapping on $\tilde{\Omega}$.) In particular, each process $(P_t(E))_{t\geq 0}$ remains probabilistically identical when considered on the extended space.

Finally, we introduce the random variable $\tau : \tilde{\Omega} \to \mathbb{E}$ as

$$\tau(\tilde{\omega}) = \tau(\omega, e) = e.$$

We then have that $P_t(\cdot)$ is indeed an $\tilde{\mathcal{F}}_t$-conditional distribution of τ, as the following result shows.

Lemma 68 *For any $E \in \mathcal{E}$, we have $P_t(E) = \tilde{P}(\tau \in E \mid \tilde{\mathcal{F}}_t)$ a.s.*

Proof. Let $\tilde{E}(\cdot)$ denote expectation under $\tilde{P}$. Since $P_t(E) = \tilde{E}(\mu(E) \mid \tilde{\mathcal{F}}_t)$, it suffices to show that $\mu(E) = \tilde{P}(\tau \in E \mid \tilde{\mathcal{F}}_\infty)$. Furthermore, $\tilde{\mathcal{F}}_\infty = \mathcal{F}_\infty \otimes \{\emptyset, \mathbb{E}\}$, so we only need to prove that

$$\tilde{E}\big(\mu(E)\mathbf{1}_{A\times B}\big) = \tilde{E}\big(\mathbf{1}_{\{\tau\in E\}}\mathbf{1}_{A\times B}\big)$$

for $A \in \mathcal{F}_\infty$ and $B \in \{\emptyset, \mathbb{E}\}$. If $B = \emptyset$ the equality is trivial, so suppose $B = \mathbb{E}$. We get

$$\begin{aligned}
\tilde{E}\big(\mu(E)\mathbf{1}_{A\times\mathbb{E}}\big) &= \int_{A\times\mathbb{E}} \mu(E;\omega) d\tilde{P}(\omega, e) \\
&= \int_A \int_{\mathbb{E}} \mu(E;\omega) d\mu(e;\omega) dP(\omega) \\
&= \int_A \mu(E;\omega) \left(\int_{\mathbb{E}} d\mu(e;\omega) \right) dP(\omega) \\
&= \int_A \mu(E;\omega) dP(\omega) \\
&= \tilde{P}(A \times E).
\end{aligned}$$

Since $\tau(\omega, e) = e$ by definition, we have $\tilde{E}(\mathbf{1}_{\{\tau \in E\}}\mathbf{1}_{A\times\mathbb{E}}) = \tilde{P}(A\times E)$. The proof is finished. ∎

What we have shown so far is that given any system of conditional distributions $P_t(\cdot)$, we can construct a random variable τ on some extended probability space such that the conditional distributions of τ are exactly $P_t(\cdot)$. Moreover, this can be done in a way that preserves the probabilistic structure of all random elements on the original space where $P_t(\cdot)$ was defined. In practice, one has to be able to simulate joint realizations of $P_t(\cdot)$ and τ, or, more generally, joint realizations of τ and any random element on Ω. Given the above construction, this amounts to simulating a realization of $\tilde{\omega} = (\omega, e)$ under the measure $\tilde{P}$. The structure of $\tilde{P}$ suggests the following recipe for doing this.

The following gives a realization of (ω, e) from $\tilde{P}$.

1. Draw ω from P, and define $\mu(\cdot) = \mu(\cdot; \omega)$.
2. Draw e from μ, independently of ω.

To the extent one finds it necessary, the correctness of the above simulation recipe can be verified as follows. Let $T : \Omega \times [0, 1] \to \mathbb{E}$ be a mapping such that for any fixed $\omega \in \Omega$, if U is uniformly distributed on $[0, 1]$, $T(\omega, U)$ is distributed according to $\mu(\cdot; \omega)$. The mapping T corresponds to the simulation routine needed to carry out step 2 of the above algorithm, so we may assume that it exists. The output from the algorithm is $(\omega, e) = (\omega, T(\omega, \omega'))$, where ω is drawn from P and ω' is drawn from Lebesgue measure λ on $[0, 1]$ independently of ω. We need to check that the distribution of (ω, e) under $P \otimes \lambda$ is precisely $\tilde{P}$. If T is measurable we get, for $A \in \mathcal{F}$ and $B \in \mathcal{E}$,

$$\begin{aligned}(P \otimes \lambda)(A \times \{\omega' : e \in B\}) &= \int_A \int_{[0,1]} \mathbf{1}_{\{T(\omega,\omega')\in B\}} d\lambda(\omega') dP(\omega) \\ &= \int_A \mu(B; \omega) dP(\omega) = \tilde{P}(A \times B),\end{aligned}$$

as desired. The assumed measurability of T is of course a pure technicality, since in practice the simulation routine operates with finite probability spaces.

It is important to emphasize that we have said nothing about how to actually carry out the two steps of the simulation algorithm. Typically Ω would be taken to be path space, in which case Step 1 consists in simulating a path from some stochastic process, for example Brownian motion. Step 2 is potentially more challenging, since it requires a procedure for simulating from any distribution in a possibly large class $\{\mu(\cdot; \omega) : \omega \in \Omega\}$. Finding suitable conditional distributions $P_t(\cdot)$ and efficient methods for simulating from them is an important problem for applications.

In the first part of this subsection it was shown how a consistent vector $\boldsymbol{\tau}$ can be constructed and simulated after conditional densities $p_t(\boldsymbol{u})$, or, equivalently, conditional distributions $P_t(\boldsymbol{u})$, have been chosen. The modeling problem thus boils down to specifying such densities. In this section we describe some possible approaches. Our aim here is merely to give a flavor of what can be done; the development of more specific and practical models, including procedures for simulation, is outside the scope of this thesis.

2.3.3.1 Multiplicative one-factor models

Let R and $\boldsymbol{U} = (U_1, \ldots, U_n)$ be independent random variables, with $R \geq 0$ and $U_i \geq 0$ for all i. The default times are given by products

$$\tau_i = RU_i, \qquad (i = 1, \ldots, n).$$

Then

$$P_t(\boldsymbol{u}) = P(\tau_1 > u_1, \ldots, \tau_n > u_n \mid \mathcal{F}_t) = \int_0^\infty P(\boldsymbol{U} > \frac{1}{r}\boldsymbol{u} \mid \mathcal{F}_t) P(R \in dr \mid \mathcal{F}_t).$$

In particular, if $\boldsymbol{U}$ is independent of $\mathcal{F}_\infty$ and R has an $\mathcal{F}_t$-conditional density $f_t(r)$ with respect to Lebesgue measure, the right side reduces to $\int_0^\infty P(\boldsymbol{U} > \frac{1}{r}\boldsymbol{u}) f_t(r) dr$.

Example 6 *Assume that $\boldsymbol{U}$ is independent of $\mathcal{F}_\infty$ and that R has an $\mathcal{F}_t$-conditional density as above. If $\boldsymbol{U}$ is uniformly distributed on the simplex $\{\boldsymbol{x} \in \mathbb{R}^n_+ : \|\boldsymbol{u}\|_1 = 1\}$, it is known that $P(\boldsymbol{U} > \boldsymbol{u}) = (1 - \|\boldsymbol{u}\|_1)_+^{n-1}$, see Theorem 5.2(i) in [47]. We get*

$$P(\tau_1 > u_1, \ldots, \tau_n > u_n \mid \mathcal{F}_t) = \int_0^\infty \left(1 - \frac{\|\boldsymbol{u}\|_1}{r}\right)_+^{n-1} f_t(r) dr.$$

Under appropriate integrability conditions, one can compute

$$\begin{aligned}\frac{\partial^{n-2} P(\tau_1 > u_1, \ldots, \tau_n > u_n \mid \mathcal{F}_t)}{\partial u_1 \ldots \partial u_{n-2}} &= (n-1)!(-1)^{n-2} \int_0^\infty \left(1 - \frac{\|\boldsymbol{u}\|_1}{r}\right)_+ r^{-n+2} f_t(r) dr \\ &= (n-1)!(-1)^{n-2} \int_{\|\boldsymbol{u}\|_1}^\infty \left(1 - \frac{\|\boldsymbol{u}\|_1}{r}\right) r^{-n+2} f_t(r) dr,\end{aligned}$$

and with some more calculation one gets

$$\frac{\partial^n P(\tau_1 > u_1, \ldots, \tau_n > u_n \mid \mathcal{F}_t)}{\partial u_1 \ldots \partial u_n} = (n-1)!(-1)^n \|\boldsymbol{u}\|_1^{-n+1} f_t(\|\boldsymbol{u}\|_1),$$

i.e.,

$$p_t(u_1, \ldots, u_n) = (n-1)! \frac{f_t(u_1 + \cdots + u_n)}{(u_1 + \cdots + u_n)^{n-1}}, \qquad \boldsymbol{u} > 0.$$

In Example 6, one could of course replace the uniform distribution on the simplex by any other distribution taking values in $\mathbb{R}^n_+$, as long as the existence of the resulting density $p_t(\boldsymbol{u})$ can be ensured. This procedure reduces the task to that of modeling the conditional density of the single random variable R.

This model can also be made amenable to simulation. Indeed, if a method for generating paths of $f_t(r)$ is available, the general scheme described in Section 2.3.3 tells us that consistent samples from $\boldsymbol{\tau}$ can be obtained by setting $\tau_i = R_T U_i$, where R_T is a draw from the density $f_T(r;\omega)$ and U_i has been generated independently (here T is some fixed time horizon.) Simulating from $f_T(r)$ is usually feasible, since it is a one-dimensional density.

2.3.3.2 Measure change approach

We now give a different way of systematically constructing $\mathcal{F}_t$-conditional densities $p_t(\boldsymbol{u})$. The starting point is to specify the dynamics of a family of non-normalized densities $q_t(\boldsymbol{u})$, and then change the measure so that each $p_t(\boldsymbol{u})$ given by

$$p_t(\boldsymbol{u}) = \frac{q_t(\boldsymbol{u})}{\int_0^\infty \cdots \int_0^\infty q_t(\boldsymbol{v})d\boldsymbol{v}}$$

becomes a martingale.

We assume that the filtration $\mathbb{F}$ is generated by a Brownian motion W. Let $q_t(\boldsymbol{u})$ be the solution to the SDE

$$dq_t(\boldsymbol{u}) = q_t(\boldsymbol{u})\sigma_t(\boldsymbol{u})dW_t,$$

where for each $\boldsymbol{u} \in \mathbb{R}^n_+$, $\sigma(\boldsymbol{u})$ is such that $q(\boldsymbol{u})$ is a strictly positive $\mathbb{F}$ martingale. Assume also that there exists an integrable process that does not depend on $\boldsymbol{u}$ that dominates every $q_t(\boldsymbol{u})\sigma_t(\boldsymbol{u})$. Define the process of normalizing constants

$$C_t = \int_0^\infty \cdots \int_0^\infty q_t(\boldsymbol{u})d\boldsymbol{u},$$

and set

$$p_t(\boldsymbol{u}) = \frac{q_t(\boldsymbol{u})}{C_t}.$$

Then $p_t(\boldsymbol{u}) \geq 0$ and $\int_0^\infty \cdots \int_0^\infty p_t(\boldsymbol{u})d\boldsymbol{u} = 1$ for all $t \geq 0$, P-a.s. The problem is that $p_t(\boldsymbol{u})$ is not a martingale. We may remedy this problem by changing to an equivalent probability Q. By the definition of $q_t(\boldsymbol{u})$ and C_t, and using the fact that $q_t(\boldsymbol{u})\sigma_t(\boldsymbol{u})$ is dominated by an integrable process, we get

$$\begin{aligned} C_t &= C_0 + \int_0^\infty \cdots \int_0^\infty \int_0^t q_s(\boldsymbol{u})\sigma_s(\boldsymbol{u})dW_s d\boldsymbol{u} \\ &= C_0 + \int_0^t \int_0^\infty \cdots \int_0^\infty q_s(\boldsymbol{u})\sigma_s(\boldsymbol{u})d\boldsymbol{u}dW_s \\ &= C_0 + \int_0^t C_s \nu_s dW_s, \end{aligned}$$

where we define

$$\nu_t = \frac{1}{C_t}\int_0^\infty \cdots \int_0^\infty q_t(\boldsymbol{u})\sigma_t(\boldsymbol{u})d\boldsymbol{u} = \int_0^\infty \cdots \int_0^\infty p_t(\boldsymbol{u})\sigma_t(\boldsymbol{u})d\boldsymbol{u}.$$

Itô's formula then implies that

$$dp_t(\boldsymbol{u}) = p_t(\boldsymbol{u})(\sigma_t(\boldsymbol{u}) - \nu_t)(dW_t - \nu_t dt).$$

Since $(C_t)_{t\geq 0}$ is a positive martingale, we may define an equivalent probability measure Q via $dQ = C_T dP$. Under Q, the process $dW_t^Q = dW_t - \nu_t dt$ is Brownian motion, and thus

$$dp_t(\boldsymbol{u}) = p_t(\boldsymbol{u})(\sigma_t(\boldsymbol{u}) - \nu_t)dW_t^Q \tag{2.8}$$

is a Q martingale. The normalization condition $\int_0^\infty \cdots \int_0^\infty p_t(\boldsymbol{u})d\boldsymbol{u} = 1$ and the nonnegativity of $p_t(\boldsymbol{u})$ hold Q-a.s., since P and Q are equivalent. Therefore, $p_t(\boldsymbol{u})$ defines a family of $\mathcal{F}_t$-conditional densities on the space $(\Omega, \mathcal{F}, \mathbb{F}, Q)$.

Notice that the quantities in (2.8) no longer depend on $q_t(\boldsymbol{u})$. Therefore, given the processes $(\sigma_t(\boldsymbol{u}))_{t\geq 0}$, we could alternatively model $p_t(\boldsymbol{u})$ directly under the original measure P as the solution of the infinite-dimensional stochastic differential equation

$$dp_t(\boldsymbol{u}) = p_t(\boldsymbol{u})\big\{\sigma_t(\boldsymbol{u}) - \int_0^\infty \cdots \int_0^\infty p_t(\boldsymbol{v})\sigma_t(\boldsymbol{v})d\boldsymbol{v}\big\}dW_t.$$

This methodology has been used by Filipović et al. [48] to model the dynamics of volatility surfaces. In fact, they show that any system of $\mathcal{F}_t$-conditional densities must have this form. Extensions of these results have also been derived by El Karoui et al. [42]. The task of finding tractable specifications and analyzing their performance on real-world data sets is important but is beyond the scope of this research.

2.3.3.3 Stochastic double-exponential

Assume that $M_t(\boldsymbol{u})$, where $\boldsymbol{u} \in \mathbb{R}_+^n$, is a continuous local $\mathbb{F}$ martingale for each $\boldsymbol{u}$. Then the processes

$$Y_t(\boldsymbol{u}) = \mathcal{E}(-\mathcal{E}(-M(\boldsymbol{u})))_t$$

are also local martingales for all $\boldsymbol{u}$. More explicitly, we have

$$\begin{aligned} Y_t(\boldsymbol{u}) &= \exp\left\{-e^{-M_t(\boldsymbol{u})-\frac{1}{2}\langle M(\boldsymbol{u})\rangle_t} - \frac{1}{2}\int_0^t e^{-2M_s(\boldsymbol{u})-\langle M(\boldsymbol{u})\rangle_s} d\langle M(\boldsymbol{u})\rangle_s\right\} \\ &= \exp\left\{-e^{-X_t(\boldsymbol{u})} - \frac{1}{2}\int_0^t e^{-2X_s(\boldsymbol{u})} d\langle X(\boldsymbol{u})\rangle_s\right\}, \end{aligned}$$

where

$$X_t(\boldsymbol{u}) = M_t(\boldsymbol{u}) + \frac{1}{2}\langle M(\boldsymbol{u})\rangle_t.$$

This shows in particular that $0 < Y_t(\boldsymbol{u}) < 1$ a.s. for all $\boldsymbol{u}$ and t. We would like to set

$$P(\boldsymbol{\tau} \leq \boldsymbol{u} \mid \mathcal{F}_t) = Y_t(\boldsymbol{u}),$$

but for this to be a valid construction, the following properties are necessary:

- $u_i \mapsto Y_t(\boldsymbol{u})$ is nondecreasing a.s. for each i and $\boldsymbol{u}_{-i}$
- $\lim_{u_i \to \infty} Y_t(\boldsymbol{u}) = 1$ a.s. for each i and $\boldsymbol{u}_{-i}$
- $\lim_{u_i \to 0} Y_t(\boldsymbol{u}) = 0$ a.s. for each i and $\boldsymbol{u}_{-i}$.

The question is therefore how to choose the processes $M(\boldsymbol{u})$ in order for these conditions to be satisfied. Notice that $Y_0(\boldsymbol{u}) = \exp\{-e^{-M_0(\boldsymbol{u})}\}$, so that we can impose a given unconditional distribution $P(\boldsymbol{\tau} \leq \boldsymbol{u})$ by requiring

$$M_0(\boldsymbol{u}) = -\ln\{-\ln P(\boldsymbol{\tau} \leq \boldsymbol{u})\}.$$

Example 7 *Let $M_t(\boldsymbol{u}) = M_0(\boldsymbol{u}) + L_t$, where $(L_t)_{t\geq 0}$ is an arbitrary continuous local $\mathbb{F}$ martingale with $L_0 = 0$, and set $M_0(\boldsymbol{u}) = -\ln\{-\ln P(\boldsymbol{\tau} \leq \boldsymbol{u})\}$. Then*

$$X_t(\boldsymbol{u}) = M_0(\boldsymbol{u}) + L_t + \frac{1}{2}\langle L\rangle_t,$$

and $u_i \mapsto X_t(\boldsymbol{u})$ is nondecreasing, $\lim_{u_i\to\infty} X_t(\boldsymbol{u}) = +\infty$ and $\lim_{u_i\to 0} X_t(\boldsymbol{u}) = -\infty$ a.s. for each i and $\boldsymbol{u}_{-i}$. Since $\langle X(\boldsymbol{u})\rangle = \langle L\rangle$ is independent of $\boldsymbol{u}$, it follows that the necessary conditions on $Y_t(\boldsymbol{u})$ are satisfied.

2.3.4 An application to risk management

The purpose of this section is to emphasize the impact of ranking the default times before expanding the filtration, as opposed to using the non-ordered times. This will be done through a simple simulation study in the case of two firms. We use the same notation as in Section 2.3.1—in particular, $\sigma_1 \leq \sigma_2$ are the ranked times, and $\widetilde{\mathbb{G}}$ is the progressive expansion of $\mathbb{F}$ with (σ_1, σ_2). For simplicity, and in order to isolate the effects we aim to highlight, we take the base filtration $\mathbb{F}$ to be the trivial filtration. In this case all conditional densities coincide with the unconditional ones. Moreover, we let the default times (τ_1, τ_2) have exponential marginal distributions with parameters λ_1 and λ_2 respectively, connected through a bivariate Gaussian copula with correlation parameter ρ. In other words,

$$P(u_1, u_2) = P(\tau_1 \leq u_1, \tau_2 \leq u_2 \mid \mathcal{F}_t) = C_\rho(e^{-\lambda_1 u_1}, e^{-\lambda_2 u_2}),$$

where the Gauss copula C_ρ is given by

$$C_\rho(x_1, x_2) = \Phi_\rho(\Phi^{-1}(x), \Phi^{-1}(y)).$$

Here $\Phi(\cdot)$ is the standard Normal distribution function, and $\Phi_\rho(\cdot,\cdot)$ is the bivariate Normal distribution function with unit variances and correlation ρ. Notice that $P(u_1, u_2)$ does not depend on t since $\mathbb{F}$ is trivial. One readily computes

$$\partial_{u_1} P(u_1, u_2) = -\lambda_1 e^{-\lambda_1 u_1} \Phi\left(\frac{\Phi^{-1}(e^{-\lambda_2 u_2}) - \rho\Phi^{-1}(e^{-\lambda_1 u_1})}{\sqrt{1-\rho^2}}\right)$$

$$\partial_{u_2} P(u_1, u_2) = -\lambda_2 e^{-\lambda_2 u_2} \Phi\left(\frac{\Phi^{-1}(e^{-\lambda_1 u_1}) - \rho\Phi^{-1}(e^{-\lambda_2 u_2})}{\sqrt{1-\rho^2}}\right).$$

Finally, the discussion in Section 2.3.3 shows that in the case where $\mathbb{F}$ is trivial, we may obtain draws from τ_1 and τ_2 simply by simulating from the unconditional joint distribution. This is straightforward: one first simulates from the copula, and then applies the inverse of the marginal distributions. With this we have all ingredients needed to simulate the quantities $P(\sigma_2 > T \mid \mathcal{G}_t)$ and $P(\sigma_2 > T \mid \widetilde{\mathcal{G}}_t)$ using the formulas in Corollary 15 and Lemma 62 in Section 2.3.1.

Figure 2.1 and Figure 2.2 contain normalized histograms based on simulations of the random variables $P(\sigma_2 > T \mid \mathcal{G}_t)$ and $P(\sigma_2 > T \mid \widetilde{\mathcal{G}}_t)$ for various values of the parameters λ_1, λ_2 and ρ. In all cases, $T = 1$ and $t = 0.5$. In a risk management context, this corresponds to simulating the distribution of the time t price of a contract that pays $\mathbf{1}_{\{\sigma_2 > T\}}$ at time T. Each figure is based on 10,000 draws. Furthermore, the histograms are based exclusively on outcomes corresponding to the event $\{\sigma_1 \leq t\} \cap \{\sigma_2 > t\}$, which is the only set on which the two random variables differ.

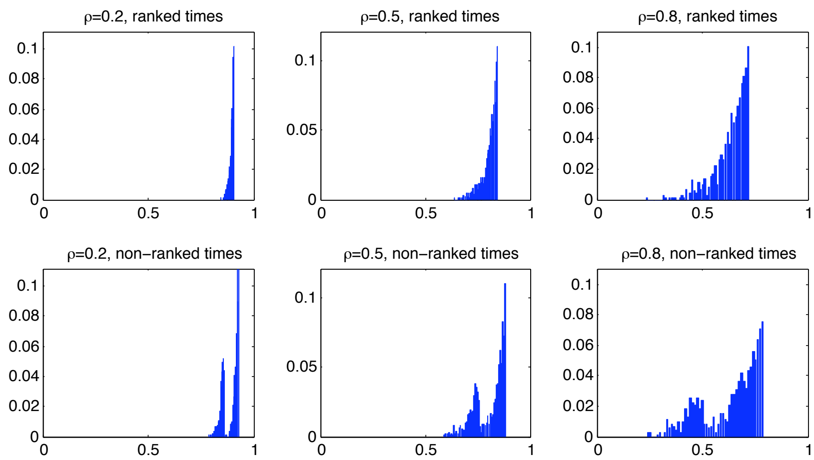

Figure 2.1: Top row: $P(\sigma_2 > T \mid \widetilde{\mathcal{G}}_t)$; Bottom row: $P(\sigma_2 > T \mid \mathcal{G}_t)$; All panels have $\lambda_1 = 0.1$, $\lambda_2 = 0.2$. ρ is as indicated above each panel.

In Figure 2.1, all six panels were generated using $\lambda_1 = 0.1$ and $\lambda_2 = 0.2$. This corresponds to unconditional marginal default probabilities equal to 0.1 for firm 1 and 0.2 for firm 2 up to the time horizon T. The left column was generated using $\rho = 0.2$, the center column using $\rho = 0.5$, and the right column $\rho = 0.8$. The top row shows $P(\sigma_2 > T \mid \widetilde{\mathcal{G}}_t)$ (ranked times), whereas the bottom row shows $P(\sigma_2 > T \mid \mathcal{G}_t)$ (non-ranked times). The figures show clearly that by ignoring the distinction between $\{\tau_1 \leq t\} \cap \{\tau_2 > t\}$ and $\{\tau_2 \leq t\} \cap \{\tau_1 > t\}$, i.e. by ranking the times, the left tail of the price distribution is heavily distorted. As the correlation parameter ρ increases, the distributions of both quantities are spread out, and, more importantly, the left peak appearing in the bottom row moves to the left. This can be expected to have significant impact on tail measures, such as value-at-risk or expected shortfall.

In Figure 2.2, the correlation parameter is fixed at $\rho = 0.5$. Also, $\lambda_1 = 0.1$, whereas λ_2

increases from 0.1 (left column) via 0.2 (center column) to 0.5 (right column). The top row again shows the case of ranked times, while the bottom row shows the case of non-ranked times. First observe that the top and bottom panels of the left column look the same. This is expected, since $\lambda_1 = \lambda_2 = 0.1$, implying that (τ_1, τ_2) and (τ_2, τ_1) have the same distribution due to the exchangeability of the bivariate Gauss copula. The center column has the same parameter values as the center column of Figure 2.1, and again the bottom panel has two separate peaks. In the bottom right panel, the smaller peak has made a significant move to the left. By contrast, the top panels all display essentially the same distribution. The main conclusion here is that an increasingly asymmetric default distribution causes the error one commits by using ranked times to become increasingly severe.

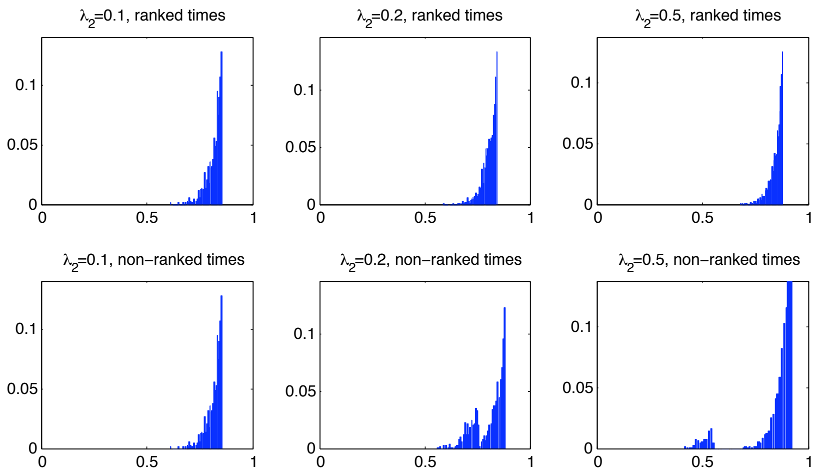

Figure 2.2: Top row: $P(\sigma_2 > T \mid \widetilde{\mathcal{G}}_t)$; Bottom row: $P(\sigma_2 > T \mid \mathcal{G}_t)$; All panels have $\rho = 0.5$, $\lambda_1 = 0.1$. λ_2 is as indicated above each panel.

In the context of reduced-form models for multiple default times, there has been an emphasis on models where defaults are assumed to be ordered. While adequate for pricing purposes, this assumption can be quite restrictive in other situations such as risk management, where the distribution of prices at future times is of interest. In this section we have so far provided a general framework capable of handling arbitrarily many non-ordered default times, under the basic assumption that conditional joint densities exist. In particular we derived default intensities and pricing formulas. The resulting expressions clearly demonstrated the presence of information-driven contagion effects, where the default of one firm instantaneously updates the market's knowledge of the state of the other firms, causing their intensities to jump. Moreover, by means of a simple simulation study we emphasized the distinction between ordered and non-ordered defaults in the risk management context. It is demonstrated that stronger dependence between defaults leads to a more pronounced difference between the ordered and non-ordered case. It is also shown that increasingly asymmetric default distributions have a similar effect. Now we focus in the remaining sections of this chapter on the information induced credit contagion effect under the conditional density assumption and then solve using the methodology developed in this section a toy example of a tractable two firms structural model.

2.3.5 The conditional density assumption and credit contagion

As mentioned earlier, the form in equation (2.6) makes it clear that the intensity of τ_k might jump upon the arrival of some other default τ_i, and hence there is an information induced contagion effect by Definition 9, unless the conditional densities have some special structure. As described in Section 2.3.1, the interpretation of this effect is that the default of firm i gives indirect information about the state of firm k, thereby causing its intensity to jump. One case where the type of contagion as in Definition 9 does not occur is when conditional independence is present, i.e. when the conditional densities factor.

2.3.5.1 Conditional independence models

Assume the default times are conditionally independent given $\mathcal{F}_t$:

$$p_t(u_1, \dots, u_n) = \prod_{i=1}^{n} p_t^i(u_i)$$

for one-dimensional densities $p_t^i(u_i)$. In this case, holding k fixed and letting $I, J \subset \{1, \dots, n\}$ satisfy $I \cup J = \{1, \dots, n\} \backslash \{k\}$ and $I \cap J = \emptyset$, (2.5) reduces simply to

$$\beta_t(u \mid I, J) = p_t^k(u).$$

This quantity is independent of I and J, meaning that it does not change depending on which of the times τ_i have already happened and which ones have not. Theorem 45 then implies that the $\mathbb{G}$ intensity of τ_k is given by

$$\lambda_t^k = \frac{p_t^k(t)}{\int_t^\infty p_t^k(u) du},$$

which clearly does not jump as other times τ_i arrive.

We study now the case where the times are linked through an archimedian copula. We prove that the only case where there is no credit contagion effect is if the copula is the independent one.

2.3.5.2 Archimedian copulas and credit contagion

Assume we are given two random times τ_1 and τ_2 finite a.s. such that

$$P(\tau_1 > 0) = P(\tau_2 > 0) = 1$$

Let $P(u_1,u_2) = P(\tau_1 > u_1, \tau_2 > u_2)$. Assume that (τ_1,τ_2) has a density p with support on $\mathbb{R}^{+*}\times\mathbb{R}^{+*}$. This condition is sufficient for τ_1 and τ_2 to have intensities λ^1 and λ^2. They are given by

$$\lambda^1_t = 1_{\{t<\tau_2\}}\frac{-\partial_{u_1}P(t,t)}{P(t,t)} + 1_{\{t\geq\tau_2\}}\frac{-\partial_{u_1,u_2}P(t,\tau_2)}{\partial_{u_2}P(t,\tau_2)}$$
$$\lambda^2_t = 1_{\{t<\tau_1\}}\frac{-\partial_{u_2}P(t,t)}{P(t,t)} + 1_{\{t\geq\tau_1\}}\frac{-\partial_{u_2,u_1}P(\tau_1,t)}{\partial_{u_1}P(\tau_1,t)}$$

We assume in the sequel that the joint density function p is continuous. In this case, there is no information induced credit contagion effect if and only if

$$\partial_{u_2}P(t,t)\partial_{u_1}P(t,t) = P(t,t)\partial_{u_1,u_2}P(t,t) \qquad \text{for all } t\geq 0 \tag{2.9}$$

Skar's theorem ensures that there exists a copula function $(x,y)\to C*x,y)$ such that

$$P(u_1,u_2) = C(P_1(u_1),P_2(u_2)) \qquad \text{for all } (u_1,u_2)$$

where $P_i(t) = P(\tau_i > t)$, $i\in\{1,2\}$. Then equation (2.9) can be re-written

$$\partial_x C(P_1(t),P_2(t))\partial_y C(P_1(t),P_2(t)) = C(P_1(t),P_2(t))\partial_{xy}C(P_1(t),P_2(t)) \text{ for all } t\geq 0 \tag{2.10}$$

We prove below that if C is an archimedian copula then there is no credit contagion effect if and only if C is the product copula, i.e. there is no dependence structure between τ_1 and τ_2.

Lemma 69 *Assume C is an archimedian copula, that is there exists a twice continuously differentiable function ψ such that*

$$C(x,y) = \psi^{-1}(\psi(x)+\psi(y)) \qquad \textit{for all } (x,y)$$

and $\psi(1)=0$, $\lim_{x\to 0} = +\infty$, $\psi'(x) < 0$ and $\psi''(x) > 0$. Then there is no information induced credit contagion (i.e. (2.10) holds) if and only if $\psi(x) = -\ln(x)$, for all $x>0$.

Proof. Basic computations give

$$\partial_x C(x,y) = \frac{\psi'(x)}{\psi'(C(x,y))} \quad \partial_y C(x,y) = \frac{\psi'(y)}{\psi'(C(x,y))} \quad \partial_{x,y}C(x,y) = \frac{-\psi'(x)\psi'(y)\psi''(C(x,y))}{\left(\psi'(C(x,y))\right)^3}$$

Therefore (2.10) can be written

$$1 = -C(P_1(t),P_2(t))\frac{\psi''(C(P_1(t),P_2(t)))}{\psi'(C(P_1(t),P_2(t)))} \qquad \text{for all } t\geq 0$$

This is equivalent to $\frac{-1}{P(t,t)} = \frac{\psi''(P(t,t))}{\psi'(P(t,t))}$ for all $t>0$. Since $f: t\to P(t,t)$ is continuous, such that $f(0)=1$ and $\lim_{x\to\infty} f(x) = 0$, f takes any value between 0 and 1, so that there is no contagion effect if and only if

$$\frac{-1}{x} = \frac{\psi''(x)}{\psi'(x)} \qquad \text{for all } 0 < x \leq 1$$

Solving this ODE and using the boundary conditions for ψ, one obtains that $\psi(x) = -\ln(x)$ for all $0 < x \leq 1$. This is the well-known product (or indenpendence) copula. ■

We provide now an example where the times do not play a symmetric role. Here, the conditional densities factor out on the half plane $u_1 \geq u_2$. There will be no credit contagion effect for firm 1 and an a.s. credit contagion effect for firm 2 as soon as the joint density of the two times is not smooth enough.

2.3.5.3 Conditional densities with a half plane factor form

Let τ_1, τ_2 two random times and $\mathbb{F}$ a given filtration. Assume the existence of $\mathcal{F}_t$-conditional densities $p_t(u_1, u_2)$

$$P(\tau_1 \in du_1, \tau_2 \in du_2 \mid \mathcal{F}_t) = p_t(u_1, u_2) du_1 du_2$$

The market filtration $\mathbb{G}$ is again the progressive expansion of $\mathbb{F}$ with (τ_1, τ_2) and we recall that in this case τ_1 and τ_2 have intensities in $\mathbb{G}$ given by

$$\lambda_t^1 = -1_{\{\tau_2 > t\}} \frac{\partial_{u_1} P_t(t,t)}{P_t(t,t)} - 1_{\{\tau_2 \leq t\}} \frac{\partial_{u_1,u_2} P_t(t,\tau_2)}{\partial_{u_2} P_t(t,\tau_2)}$$

and

$$\lambda_t^2 = -1_{\{\tau_1 > t\}} \frac{\partial_{u_2} P_t(t,t)}{P_t(t,t)} - 1_{\{\tau_1 \leq t\}} \frac{\partial_{u_2,u_1} P_t(\tau_1,t)}{\partial_{u_1} P_t(\tau_1,t)}$$

We assume in this subsection that the survival probabilities do have a particular form : they factor out on the quadrant $\{u_2 \leq u_1\}$. The following holds.

Lemma 70 *Assume the survival probability P has the following form*

$$P_t(u_1, u_2) = P_t^1(u_1) P_t^2(u_2) \qquad \textit{for all } u_2 \leq u_1.$$

for one-dimensional survival probabilities $P_t^i(u_i)$ that are continuously differentiable. Assume that for each $t \geq 0$, the function $u_1 \to \frac{\partial}{\partial u_1} P_t(u_1, t)$ is continuous in t. Then

(i) λ^1 does not jump at the arrival of the default time τ_2.

(ii) λ^2 does not jump at τ_1 a.s. if and only if $\frac{p_t(\tau_1,t)}{\int_t^\infty p_t(\tau_1,v_2)dv_2}$ is continuous in probability at τ_1.

(iii) Furthermore, if for each u_1 in the support of τ_1, the function $t \to \frac{\partial}{\partial u_1} P_t(u_1, t)$ is continuous in u_1 then λ^2 does not jump at τ_1 a.s. if and only if $p_t(\tau_1, t)$ is continuous in probability at τ_1.

Proof. We start with (i). Using that $P_t(u_1, u_2) = P_t^1(u_1) P_t^2(u_2)$ for $u_2 \leq u_1$, we get, on $\{\tau_2 \leq t\}$,

$$\frac{\partial_{u_1,u_2} P_t(t,\tau_2)}{\partial_{u_2} P_t(t,\tau_2)} = \frac{\partial_u P_t^1(t)}{P_t^1(t)}$$

Now for $t \leq u_1$, $\frac{\partial}{\partial u_1} P_t(u_1, t) = P_t^2(t) \frac{\partial P_t^1}{\partial u}(u_1) \to P_t^2(t) \frac{\partial P_t^1}{\partial u}(t)$ when $u_1 \to t^-$. Since $u_1 \to \frac{\partial}{\partial u_1} P_t(u_1, t)$ is continuous in t,

$$\frac{1}{P_t(t,t)} \frac{\partial P_t(t,t)}{\partial u_1} = \frac{\partial_u P_t^1(t)}{P_t^1(t)}$$

and it is clear from the expressions of the intensities that λ^1 does not jump at τ_2. We prove now (ii). First notice that for $t \geq u_1$,

$$\frac{\partial}{\partial u_1} P_t(u_1, t) = -\int_t^\infty p_t(u_1, v_2) dv_2$$

so by continuity in t of this function,

$$-\int_t^\infty p_t(t, v_2) dv_2 = P_t^2(t) \frac{\partial P_t^1}{\partial_u}(t) \tag{2.11}$$

Using again the factorized structure of P, we obtain for $t \leq u_1$, $\frac{\partial_{u_2} P_t(u_1,t)}{P_t(u_1,t)} \to \frac{\partial_u P_t^2(t)}{P_t^2(t)}$ when $u_1 \to t^+$. For $u_2 \geq u_1$, notice that

$$\begin{aligned}
P_t(u_1, u_2) &= P(\tau_1 > u_1, \tau_2 > u_2 \mid \mathcal{F}_t) \\
&= P(\tau_1 > u_2, \tau_2 > u_2 \mid \mathcal{F}_t) + P(u_1 < \tau_1 \leq u_2, \tau_2 > u_2 \mid \mathcal{F}_t) \\
&= P_t^1(u_2) P_t^2(u_2) + \int_{u_1}^{u_2} \int_{u_2}^\infty p_t(v_1, v_2) dv_2 dv_1
\end{aligned}$$

to deduce that for $t \geq u_1$,

$$\frac{\partial_{u_2} P_t(u_1, t)}{P_t(u_1, t)} = \frac{1}{P_t(u_1, t)} \Big(P_t^2(t) \frac{\partial P_t^1}{\partial u}(t) + P_t^1(t) \frac{\partial P_t^2}{\partial u}(t) - \int_{u_1}^t p_t(v_1, t) dv_1 + \int_t^\infty p_t(u_1, v_2) dv_2 \Big)$$

This quantity converges to

$$\frac{\partial_u P_t^1(t)}{P_t^1(t)} + \frac{\partial_u P_t^2(t)}{P_t^2(t)} + \frac{1}{P_t^1(t) P_t^2(t)} \int_t^\infty p_t(t, v_2) dv_2$$

when $u_1 \to t^-$. Using equation (2.11) we obtain that this limit is equal to $\frac{\partial_u P_t^2(t)}{P_t^2(t)}$ and $u_1 \to \frac{\partial_{u_2} P_t(u_1,t)}{P_t(u_1,t)}$ is continuous in t. Finally

$$\frac{\partial_{u_2} P_t(t,t)}{P_t(t,t)} = \frac{\partial_u P_t^2(t)}{P_t^2(t)}$$

Now, on $\{\tau_1 \leq t\}$,

$$\frac{\partial_{u_2, u_1} P_t(\tau_1, t)}{\partial_{u_1} P_t(\tau_1, t)} = -\frac{p_t(\tau_1, t)}{\int_t^\infty p_t(\tau_1, v_2) dv_2}$$

Using the expression for λ^2, there is no contagion effect at the default time τ_1 if and only if

$$P\Big(\lim_{t \to \tau_1^-} \frac{\partial_u P_t^2(t)}{P_t^2(t)} = -\frac{p_t(\tau_1, t)}{\int_t^\infty p_t(\tau_1, v_2) dv_2} \Big|_{t=\tau_1} \Big) = 1$$

It follows from equation (2.11) that

$$\lim_{t\to\tau_1^-}\frac{\partial_u P_t^2(t)}{P_t^2(t)} = -\lim_{t\to\tau_1^-}\frac{\partial_u P_t^2(t)\partial_u P_t^1(t)}{\int_t^\infty p_t(t,v_2)dv_2}$$

So λ^2 does not jump at τ_1 if and only if

$$P(\lim_{t\to\tau_1^-}\frac{\partial_u P_t^2(t)\partial_u P_t^1(t)}{\int_t^\infty p_t(t,v_2)dv_2} = \frac{p_t(\tau_1,t)}{\int_t^\infty p_t(\tau_1,v_2)dv_2}\Big|_{t=\tau_1}) = 1$$

Since $t\to\partial_u P_t^1(u_1)$ is continuous in u_1 and since $p_t(u_1,t)=\partial_u P_t^2(t)\partial_u P_t^1(u_1)$ for each $t\le u_1$, this is equivalent to

$$P(\lim_{t\to\tau_1^-}\frac{p_t(\tau_1,t)}{\int_t^\infty p_t(t,v_2)dv_2} = \frac{p_t(\tau_1,t)}{\int_t^\infty p_t(\tau_1,v_2)dv_2}\Big|_{t=\tau_1}) = 1$$

which yields (ii). Point (iii) follows from (ii) by noticing that $\int_t^\infty p_t(\tau_1,v_2)dv_2 = \frac{\partial}{\partial u_1}P_t(u_1,t)$ is continuous in τ_1. ∎

We link now explicitly this section and the previous one on the information induced credit contagion effect in structural models, and suggest a toy structural model, which satisfies the assumptions of Lemma 70 and compute explicitly the jump size of the intensity of τ_2 at the default time τ_1. We pointed out many times now that the arrival of a default time will usually help to localize the other default time in the support of its conditional density which we already interpreted as the default of the first firm indirectly giving information about the state of the second firm, and vice versa. This is exactly what happens for the second firm in our example at the default time of the first firm. However, the structure of the structural model below is such that there is no credit contagion effect for the first firm at the default of the second one. This is due to the factor form that takes the joint density when $u_1\ge u_2$.

2.3.6 A toy structural model and credit contagion

Let $(\Omega,\mathcal{F},P,\mathbb{F})$ be a given filtered probability space that supports two independent Brownian motions B^1 and B^2. Consider the model

$$dX_t^1 = dB_t^1$$
$$dX_t^2 = dB_t^2 + f(X_{t\wedge\tau_1}^1)dt$$

where $\tau_i=\inf\{t\mid X_t^i\le\alpha_i\}$, where $\alpha_1<0$ and $\alpha_2<0$ are given and $f(x)=f_1 1_{\{x>\alpha_1\}}+f_2 1_{\{x\le\alpha_1\}}$, with f_1 and f_2 two constants. Assume that $X_0^1=X_0^2=0$.

Introduce the running minimum processes $m_t^i=\inf_{s\le t}X_s^i$, and for a given generic Brownian motion B and a constant μ, define $m_t^{B,\mu}=\inf_{s\le t}B_s+\mu s$. Finally, introduce the function $G(t,\mu,\alpha)=P(m_t^{\mu,B}>\alpha)$, which does not depend on the Brownian motion B. The following lemma is very classical.

Lemma 71
$$G(t,\mu,\alpha) = \mathcal{N}(-\frac{\alpha}{\sqrt{t}} + \mu\sqrt{t}) - e^{2\mu\alpha}\mathcal{N}(\frac{\alpha}{\sqrt{t}} + \mu\sqrt{t})$$
In particular, $G(t,0,\alpha) = 1 - 2\mathcal{N}(\frac{\alpha}{\sqrt{t}})$.

Proof. Let Q be the probability measure equivalent to P defined through its Radon-Nykodim derivative $\frac{dQ}{dP} = e^{-\mu B_t - \frac{1}{2}\mu^2 t}$. Under Q, $W_t := B_t + \mu t$ is a Brownian motion and

$$\begin{aligned} G(t,\mu,\alpha) &= E^Q(e^{\mu W_t - \frac{1}{2}\mu^2 t} 1_{\{\inf_{s\le t} W_s > \alpha\}}) \\ &= E^Q(e^{\mu W_t - \frac{1}{2}\mu^2 t}(1 - e^{-2\frac{-\alpha(-\alpha+W_t)}{t}}) 1_{\{W_t > \alpha\}}) \end{aligned}$$

by conditioning w.r.t W_t and using $P(m_t^{W,0} > \alpha \mid W_t = x) = 1 - e^{-\frac{2\alpha(\alpha - x)}{t}}$. Continuying the computation, we obtain

$$\begin{aligned} G(t,\mu,\alpha) &= \int_{\frac{\alpha}{\sqrt{t}}}^{\infty} e^{\mu\sqrt{t}y - \frac{1}{2}\mu^2 t}(1 - e^{2\frac{\alpha(\sqrt{t}y - \alpha)}{t}})\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \\ &= \frac{1}{\sqrt{2\pi}}\Big(\int_{\frac{\alpha}{\sqrt{t}}}^{\infty} e^{-\frac{1}{2}(y-\mu\sqrt{t})^2} dy - e^{2\mu\alpha}\int_{\frac{\alpha}{\sqrt{t}}}^{\infty} e^{-\frac{1}{2}(y-(\mu\sqrt{t}+2\frac{\alpha}{\sqrt{t}}))^2} dy\Big) \\ &= \mathcal{N}(-\frac{\alpha}{\sqrt{t}} + \mu\sqrt{t}) - e^{2\mu\alpha}\mathcal{N}(\frac{\alpha}{\sqrt{t}} + \mu\sqrt{t}) \end{aligned}$$

■

For any $u_1 > 0$ and $u_2 > 0$, define $P(u_1,u_2) = P(\tau_1 > u_1, \tau_2 > u_2)$. We compute $P(u_1,u_2)$ in the two cases $u_2 \le u_1$ and $u_2 < u_1$.

Lemma 72 *If* $u_2 \le u_1$*, then*

$$P(u_1,u_2) = G(u_1,0,\alpha_1)G(u_2,f_1,\alpha_2)$$

Proof. On $\{\tau_1 > u_1\}$, for all $t \le u_1$, $X_t^1 > \alpha_1$ a.s and $f(X_t^1) = f_1$ a.s. Therefore $X_t^2 = B_t^2 + f_1 t$ for all $t \le u_1$ on $\{\tau_1 > u_1\}$. Let $dX_t = dB_t^2 + f_1 dt$. Then on $\{\tau_1 > u_1\}$,

$$\{\tau_2 > u_2\} = \{m_{u_2}^2 > \alpha_2\} = \{m_{u_2}^{B^2,f_1} > \alpha_2\}$$

Therefore, by the independence of B^1 and B^2,

$$P(\tau_1 > u_1, \tau_2 > u_2) = E(1_{\{m_{u_1}^1 > \alpha_1\}} 1_{\{m_{u_2}^{B^2,f_1} > \alpha_2\}}) = P(m_{u_1}^1 > \alpha_1)P(m_{u_2}^{B^2,f_1} > \alpha_2)$$

We can conclude using Lemma 71 ■

We focus now on the case $u_2 > u_1$.

Lemma 73 *If* $u_2 > u_1$*, then*

$$\begin{aligned} P(u_1,u_2) &= G(u_2,0,\alpha_1)G(u_2,f_1,\alpha_2) \\ &+ \int_{u_1}^{u_2}\Big(\int_{\alpha_2}^{\infty} e^{f_1 x - \frac{1}{2}f_1^2 t}(1 - e^{-\frac{2\alpha_2(\alpha_2 - x)}{t}})G(u_2 - t, f_2, \alpha_2 - x)\frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} dx\Big) g_1(t) dt \end{aligned}$$

where $g_1(t) = \frac{\alpha_1}{\sqrt{2\pi t^3}} e^{-\frac{\alpha_1^2}{2t}}$ *is the density of* τ_1.

Proof. First notice that

$$P(\tau_1 > u_1, \tau_2 > u_2) = P(\tau_1 > u_2, \tau_2 > u_2) + P(u_1 < \tau_1 \leq u_2, \tau_2 > u_2)$$

Using Lemma 72,

$$P(\tau_1 > u_2, \tau_2 > u_2) = G(u_2, 0, \alpha_1) G(u_2, f_1, \alpha_2)$$

so it only remains to compute $P(u_1 < \tau_1 \leq u_2, \tau_2 > u_2)$. On $\{\tau_1 \leq u_2\}$,

$$\{\tau_2 > u_2\} = \{\inf_{s \leq \tau_1} B_s^2 + f_1 s > \alpha_2, \inf_{\tau_1 \leq s < \tau_2} B_s^2 + f_1\tau_1 + f_2(s - \tau_1) > \alpha_2\}$$

Let $m_t^{B,u,\mu} = \inf_{t \leq s \leq u} B_s + \mu s$. Then

$$\begin{aligned}
P(u_1 < \tau_1 \leq u_2, \tau_2 > u_2) &= P(u_1 < \tau_1 \leq u_2, m_{\tau_1}^{B^2,f_1} > \alpha_2, m_{\tau_1}^{B^2,u_2,f_2} > \alpha_2 + (f_2 - f_1)\tau_1) \\
&= \int_{u_1}^{u_2} g_1(t) P(m_t^{B^2,f_1} > \alpha_2, m_t^{B^2,u_2,f_2} > \alpha_2 + (f_2 - f_1)t) dt \\
&= \int_{u_1}^{u_2} g_1(t) E\Big(1_{\{m_t^{B^2,f_1} > \alpha_2\}} P\big(m_t^{B^2,u_2,f_2} > \alpha_2 + (f_2 - f_1)t \mid \mathcal{F}_t^2\big)\Big) dt
\end{aligned}$$

We compute the conditional probability

$$\begin{aligned}
P\big(m_t^{B^2,u_2,f_2} > x \mid \mathcal{F}_t^2\big) &= P(\inf_{t \leq s \leq u_2} B_s^2 - B_t^2 + f_2(s - t) > x - f_2 t - B_t^2 \mid \mathcal{F}_t^2) \\
&= P\Big(\inf_{u \leq u_2 - t} \tilde{B}_u + f_2 u > x - f_2 t - b\Big)\Big|_{b = B_t^2} = P\Big(m_{u_2 - t}^{\tilde{B},f_2} > x - (f_2 t + b)\Big)\Big|_{b = B_t^2} \\
&= G\big(u_2 - t, f_2, x - (f_2 t + B_t^2)\big)
\end{aligned}$$

where $\tilde{B}$ is a generic Brownian motion. Therefore

$$\begin{aligned}
P(u_1 < \tau_1 \leq u_2, \tau_2 > u_2) &= \int_{u_1}^{u_2} g_1(t) E\Big(1_{\{m_t^{B^2,f_1} > \alpha_2\}} G\big(u_2 - t, f_2, \alpha_2 - (f_1 t + B_t^2)\big)\Big) dt \\
&= \int_{u_1}^{u_2} g_1(t) E^Q\Big(e^{f_1 \tilde{B}_t^2 - \frac{1}{2} f_1^2 t} 1_{\{\inf_{s \leq t} \tilde{B}_s^2 > \alpha_2\}} G\big(u_2 - t, f_2, \alpha_2 - \tilde{B}_t^2\big)\Big) dt
\end{aligned}$$

where Q is a probability measure equivalent to P with $\frac{dQ}{dP} = e^{-f_1 B_t^2 - \frac{1}{2} f_1^2 t}$ and $\tilde{B}_t^2 = B_t^2 + f_1 t$ is a Q Brownian motion. Finally, conditioning w.r.t $\tilde{B}_t^2$ and using

$$P(m_t^{\tilde{B}^2,0} > \alpha_2 \mid \tilde{B}_t^2 = x) = 1 - e^{-\frac{2\alpha_2(\alpha_2 - x)}{t}}$$

we obtain that

$$\begin{aligned}
&P(u_1 < \tau_1 \leq u_2, \tau_2 > u_2) = \\
&\quad \int_{u_1}^{u_2} g_1(t) \int_{\alpha_2}^{\infty} e^{f_1 x - \frac{1}{2} f_1^2 t} G(u_2 - t, f_2, \alpha_2 - x)(1 - e^{-\frac{2\alpha_2(\alpha_2 - x)}{t}}) \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} dx dt
\end{aligned}$$

That $g_1(t) = \frac{\alpha_1}{\sqrt{2\pi t^3}} e^{-\frac{\alpha_1^2}{2t}}$ is a well known fact. Putting everything together yields the result.
■

We can now study the credit contagion effect. $P(u_1, u_2)$ is clearly differentiable and τ_1 and τ_2 admit a density. We know that τ_2 has an intensity λ^2, given by

$$\lambda_t^2 = 1_{\{t<\tau_1\}} \frac{-\partial_{u_2} P(t,t)}{P(t,t)} + 1_{\{t\geq\tau_1\}} \frac{-\partial_{u_2,u_1} P(\tau_1,t)}{\partial_{u_1} P(\tau_1,t)}$$

and τ_1 has an intensity λ^1 given by

$$\lambda_t^1 = 1_{\{t<\tau_2\}} \frac{-\partial_{u_1} P(t,t)}{P(t,t)} + 1_{\{t\geq\tau_2\}} \frac{-\partial_{u_1,u_2} P(t,\tau_2)}{\partial_{u_2} P(t,\tau_2)}$$

as in equation (2.6). The term $\frac{\partial_{u_2} P(t,t)}{P(t,t)}$ can be computed explicitly using Lemma 72 while the term $\frac{\partial_{u_2,u_1} P(\tau_1,t)}{\partial_{u_1} P(\tau_1,t)}$ can be computed explicitly using Lemma 73 since we are working on the set $\{t \geq \tau_1\}$. The quantities appearing in the expression for λ^1 can also be computed using these two lemmas. It only remains to check if

$$P\Big(\lim_{t\to\tau_1^-} \frac{\partial_{u_2} P(t,t)}{P(t,t)} \neq \frac{\partial_{u_2,u_1} P(\tau_1,\tau_1)}{\partial_{u_1} P(\tau_1,\tau_1)}\Big) > 0$$

and/or

$$P\Big(\lim_{t\to\tau_2^-} \frac{\partial_{u_1} P(t,t)}{P(t,t)} \neq \frac{\partial_{u_1,u_2} P(\tau_2,\tau_2)}{\partial_{u_2} P(\tau_2,\tau_2)}\Big) > 0$$

Lemma 74 *There is no credit contagion effect for firm 1 at the default time of firm 2, i.e. λ^1 does not jump at the default time τ_2.*

Proof. Let $g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. On $\{\tau_2 \leq t\}$, from Lemma 72

$$-\frac{\partial_{u_1,u_2} P(t,\tau_2)}{\partial_{u_2} P(t,\tau_2)} = -\frac{\frac{\partial}{\partial t} G(t,0,\alpha_1)}{G(t,0,\alpha_1)} = \frac{-\alpha_1 g(\frac{\alpha_1}{\sqrt{t}})}{t^{\frac{3}{2}}(1-2\mathcal{N}(\frac{\alpha_1}{\sqrt{t}}))}$$

We prove now that $u_1 \to \frac{\partial}{\partial u_1} P(u_1,t)$ is continuous in t, for all $t \geq 0$ and compute $-\frac{\partial_{u_1} P(t,t)}{P(t,t)}$. It follows from Lemma 72 that

$$\lim_{u_1\to t^+} -\frac{\partial_{u_1} P(u_1,t)}{P(u_1,t)} = \frac{\frac{\partial}{\partial t} G(t,0,\alpha_1)}{G(t,0,\alpha_1)}$$

Now from the formula in Lemma 73

$$\begin{aligned}
\lim_{u_1\to t^-} -\frac{\partial_{u_1} P(u_1,t)}{P(u_1,t)} &= -\frac{g_1(t)\int_{\alpha_2}^{\infty} e^{f_1 x - \frac{1}{2} f_1^2 t}(1-e^{-2\frac{\alpha_2(\alpha_2-x)}{t}}) G(0,f_2,\alpha_2-x)\frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}\,dx}{G(t,0,\alpha_1)G(t,f_1,\alpha_2)} \\
&= -\frac{\alpha_1}{t^{\frac{3}{2}}} g(\frac{\alpha_1}{\sqrt{t}}) \frac{\int_{\alpha_2}^{\infty} e^{f_1 x - \frac{1}{2} f_1^2 t}(1-e^{-2\frac{\alpha_2(\alpha_2-x)}{t}})\frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}\,dx}{G(t,0,\alpha_1)G(t,f_1,\alpha_2)} \\
&= \frac{-\alpha_1 g(\frac{\alpha_1}{\sqrt{t}})}{t^{\frac{3}{2}}(1-2\mathcal{N}(\frac{\alpha_1}{\sqrt{t}}))} = \frac{\frac{\partial}{\partial t} G(t,0,\alpha_1)}{G(t,0,\alpha_1)}
\end{aligned}$$

since the numerator in the second line of the display above equals $G(t, f_1, \alpha_2)$. This proves that

$$-\frac{\partial_{u_1} P(t,t)}{P(t,t)} = \frac{-\alpha_1 g(\frac{\alpha_1}{\sqrt{t}})}{t^{\frac{3}{2}}(1-2\mathcal{N}(\frac{\alpha_1}{\sqrt{t}}))} = -\frac{\partial_{u_1,u_2} P(t,\tau_2)}{\partial_{u_2} P(t,\tau_2)}$$

Finally, there is no contagion effect with probability one. ∎

In the following lemma, we assume for simplicity of the computations that f_2 is zero identically. We expect that there is a credit contagion effect if and only if $f_1 \neq 0$. This turns out to be the case.

Lemma 75 *There is a credit contagion effect for firm 2 at the default time of firm 1 if and only if* $f_1 \neq 0$.

Proof. On $\{\tau_1 \leq t\}$, we can use Lemma 73 to compute $\partial_{u_2,u_1} P(\tau_1, t)$ and $\partial_{u_1} P(\tau_1, t)$. First,

$$\begin{aligned}\partial_{u_1} P(\tau_1,t) &= -g_1(\tau_1)\int_{\alpha_2}^{\infty} e^{f_1 x - \frac{1}{2}f_1^2\tau_1}(1-e^{-2\frac{\alpha_2(\alpha_2-x)}{\tau_1}})G(t-\tau_1,0,\alpha_2-x)\frac{1}{\sqrt{2\pi\tau_1}}e^{-\frac{x^2}{2\tau_1}}dx \\ &= -g_1(\tau_1)\int_{\alpha_2}^{\infty} e^{f_1 x - \frac{1}{2}f_1^2\tau_1}(1-e^{-2\frac{\alpha_2(\alpha_2-x)}{\tau_1}})(1-2\mathcal{N}(\frac{\alpha_2-x}{\sqrt{t-\tau_1}}))\frac{1}{\sqrt{2\pi\tau_1}}e^{-\frac{x^2}{2\tau_1}}dx\end{aligned}$$

Using the Monotone Convergence Theorem, this converges to

$$-g_1(\tau_1)\int_{\alpha_2}^{\infty} e^{f_1 x - \frac{1}{2}f_1^2\tau_1}(1-e^{-2\frac{\alpha_2(\alpha_2-x)}{\tau_1}})\frac{1}{\sqrt{2\pi\tau_1}}e^{-\frac{x^2}{2\tau_1}}dx = -g_1(\tau_1)G(\tau_1,f_1,\alpha_2)$$

when $t \to \tau_1^+$. We turn now to $\lim_{t\to\tau_1^+} \partial_{u_2,u_1} P(\tau_1,t)$. For $u_2 \geq u_1$, differentiating

$$\partial_{u_1} P(u_1,u_2) = -g_1(u_1)\Big(G(u_1,f_1,\alpha_2) - \frac{2}{\sqrt{\pi}}e^{-\frac{1}{2}c_1^2}\int_{-\infty}^{0}(e^{c_1 z}-e^{c_2 z})e^{-\frac{z^2}{2}}\mathcal{N}(\frac{z\sqrt{u_1}}{\sqrt{u_2-u_1}})dz\Big)$$

where $c_1 = -f_1\sqrt{u_1} + \frac{\alpha_2}{\sqrt{u_1}}$ and $c_2 = -(f_1\sqrt{u_1} + \frac{\alpha_2}{\sqrt{u_1}})$ we obtain

$$\partial_{u_2,u_1} P(u_1,u_2) = \frac{2}{\sqrt{2\pi}}g_1(u_1)e^{-\frac{1}{2}c_1^2}\int_{-\infty}^{0}(e^{c_1 z}-e^{c_2 z})e^{-\frac{z^2}{2}}\frac{-z\sqrt{u_1}}{2(u_2-u_1)^{\frac{3}{2}}}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\frac{z^2 u_1}{u_2-u_1}}dz$$

Denoting $\gamma = \sqrt{\frac{u_2}{u_2-u_1}}$ and using the identity

$$\int_{-\infty}^{0} z e^{-\frac{1}{2}z^2\gamma^2 + c_i z}dz = -\frac{1}{\gamma^2} + \frac{c_i}{\gamma^3}e^{\frac{c_i^2}{2\gamma^2}}\sqrt{2\pi}\mathcal{N}(-\frac{c_i}{\gamma})$$

we obtain

$$\begin{aligned}\partial_{u_2,u_1} P(u_1,u_2) &= \frac{1}{\sqrt{2\pi}}g_1(u_1)e^{-\frac{1}{2}c_1^2}\frac{-\sqrt{u_1}}{u_2^{\frac{3}{2}}} \\ &\Big(c_1 e^{\frac{c_1^2(u_2-u_1)}{u_2}}\mathcal{N}(-\frac{c_1(u_2-u_1)}{u_2}) + c_2 e^{\frac{c_2^2(u_2-u_1)}{u_2}}\mathcal{N}(-\frac{c_2(u_2-u_1)}{u_2})\Big)\end{aligned}$$

Therefore

$$\lim_{t\to\tau_1^+}\partial_{u_2,u_1}P(\tau_1,t)=\frac{1}{\sqrt{2\pi}}\frac{\alpha_2 g_1(\tau_1)}{\tau_1^{\frac{3}{2}}}e^{-\frac{1}{2}(\frac{\alpha_2}{\sqrt{\tau_1}}-f_1\sqrt{\tau_1})^2}$$

Finally

$$\lim_{t\to\tau_1^+}-\frac{\partial_{u_2,u_1}P(\tau_1,t)}{\partial_{u_1}P(\tau_1,t)}=\frac{1}{\sqrt{2\pi}}\frac{\alpha_2}{\tau_1^{\frac{3}{2}}}\frac{e^{-\frac{1}{2}(\frac{\alpha_2}{\sqrt{\tau_1}}-f_1\sqrt{\tau_1})^2}}{G(\tau_1,f_1,\alpha_2)}$$

Using continuity arguments, it is easy to check as in the previous lemma that

$$\lim_{t\to\tau_1^-}-\frac{\partial_{u_2}P(t,t)}{P(t,t)}=-\frac{\frac{\partial}{\partial t}G(\tau_1,f_1,\alpha_2)}{G(\tau_1,f_1,\alpha_2)}$$

Recalling that

$$\frac{\partial G}{\partial t}(t,f,\alpha)=\frac{1}{2\sqrt{2\pi}}\Big(\big(\frac{\alpha}{t\sqrt{t}}+\frac{f}{\sqrt{t}}\big)e^{-\frac{1}{2}(-\frac{\alpha}{\sqrt{t}}+f\sqrt{t})^2}-e^{2f\alpha}\big(\frac{-\alpha}{t\sqrt{t}}+\frac{f}{\sqrt{t}}\big)e^{-\frac{1}{2}(\frac{\alpha}{\sqrt{t}}+f\sqrt{t})^2}\Big)$$

the default intensity λ^2 jumps at the default time τ_1 with positive probability if and only if $f_1\neq 0$. In this case, the contagion effect can be quantified and the jump size at the default time τ_1 is given by

$$\frac{1}{G(\tau_1,f_1,\alpha_2)}\big(\frac{1}{\sqrt{2\pi}}\frac{\alpha_2}{\tau_1^{\frac{3}{2}}}e^{-\frac{1}{2}(\frac{\alpha_2}{\sqrt{\tau_1}}-f_1\sqrt{\tau_1})^2}+\frac{\partial}{\partial t}G(\tau_1,f_1,\alpha_2)\big)$$

■

Note that in this example, the density of (τ_1,τ_2) is not continuous in points (t,t), in the sense that for each $u_1>0$, the function $u_2\to p(u_1,u_2)$ is discontinuous at u_1.

Lemma 76 *For each $t>0$, the function $u_2\to p(t,u_2)$ is discontinuous at t if and only if $f_1\neq 0$.*

Proof. Recall that we are assuming that f_2 is zero. For any $u_2<u_1$,

$$p(u_1,u_2)=\partial_u G(u_1,0,\alpha_1)\partial_u G(u_2,f_1,\alpha_2)=-\frac{\alpha_1 g(\frac{\alpha_1}{\sqrt{u_1}})}{u_1^{\frac{3}{2}}}\partial_u G(u_2,f_1,\alpha_2)$$

so that $\lim_{u_2\to t^+}p(t,u_2)=-\frac{\alpha_1 g(\frac{\alpha_1}{\sqrt{t}})}{t^{\frac{3}{2}}}\partial_u G(t,f_1,\alpha_2)$. For any $u_2>u_1$, using the computations from the previous lemma,

$$\lim_{u_2\to t^+}p(t,u_2)=\frac{1}{\sqrt{2\pi}}\frac{\alpha_2 g_1(t)}{t^{\frac{3}{2}}}e^{-\frac{1}{2}(\frac{\alpha_2}{\sqrt{t}}-f_1\sqrt{t})^2}=-\frac{\alpha_1 g(\frac{\alpha_1}{\sqrt{u_1}})}{u_1^{\frac{3}{2}}}\big(\frac{-\alpha_2 e^{-\frac{1}{2}(\frac{\alpha_2}{\sqrt{t}}-f_1\sqrt{t})^2}}{\sqrt{2\pi}t^{\frac{3}{2}}}\big)$$

There is then a jump in the density given by

$$-\frac{\alpha_1 g(\frac{\alpha_1}{\sqrt{u_1}})}{u_1^{\frac{3}{2}}}\Big(\partial_u G(t,f_1,\alpha_2)-\big(\frac{-\alpha_2 e^{-\frac{1}{2}(\frac{\alpha_2}{\sqrt{t}}-f_1\sqrt{t})^2}}{\sqrt{2\pi}t^{\frac{3}{2}}}\big)\Big)$$

unless $f_1 = 0$. ■

In this chapter, we first looked into different structural models under different sets of partial information and studied the information induced credit contagion effect. We achieved first this study through examples from different classes of structural models, that are usually academically accepted as standard models. We computed as explicitly as we were able to do the default intensities. They usually have very untractable forms, but seem to jump at the defaults of the other firms, unless the models have some very particular structures, mainly if they exhibit some kind of conditional independence property. Then, since in practice many structural models can fall within the large class of models where the default times admit a conditional density (conditionally to the base filtration), we focused our efforts in the second section on times satisfying this assumption and we provided partial results that allow to reconcile multiple firms structural models where the times admit a conditional density with reduced form models from the credit contagion effect's perspective. In order to do this, we slightly extended the usual filtration expansion approach with random times to the case where there is an arbitrary number of default times under the conditional density assumption and carried out the analysis without imposing any restrictions on the ordering of the individual times. We pointed out *en passant* how inadequate it can be to use the ranked versions of the times and gave an example from risk management, where the distribution of securities prices at some future time is of interest. Finally, we studied the link between credit contagion and some explicit forms of the conditional densities of the default times, and studied in detail the credit contagion effect within a toy structural model using our methodology.

ACKNOWLEDGMENT.

Parts of this chapter are based on a joint work (see [94]) with Martin Larsson.

CHAPTER 3

Bubbles: martingale theory and real time detection

Contents

3.1 Introduction

We are interested in this chapter in the detection of financial bubbles. This question is a timely one. Recently William Dudley, the President of the New York Federal Reserve, in an interview with Planet Money [58] stated "...what I am proposing is that we try to identify bubbles in real time, try to develop tools to address those bubbles, try to use those tools when appropriate to limit the size of those bubbles and, therefore, try to limit the damage when those bubbles burst." It is also widely recognized that this is not an easy task. Indeed, in 2009 the Federal Reserve Chairman Ben Bernanke said in Congressional Testimony [12] "it is extraordinarily difficult in real time to know if an asset price is appropriate or not." Without a quantitative procedure, experts often have different opinions about the existence of price bubbles. A famous example is the oil price bubble of 2007/2008. Nobel prize winning economist Paul Krugman wrote in the New York Times that it was not a bubble, and two days later Ben Stein wrote in the same paper that it was.

Asset price bubbles have also been recently characterized in frictionless, competitive, and continuous trading economies using the abitrage-free martingale pricing technology underlying option pricing theory (see [107], [108], [31], [61], [83] and [84]). In this classical setting, Jarrow, Protter and Shimbo [83] and [84] show there are three types of asset price bubbles possible. Two of these price bubbles exist only in infinite horizon economies, the third - called type 3 bubbles - exist in finite horizon settings. Consequently, type 3 bubbles are those most relevant to actual market experiences. The martingale theory of bubbles as developed by Jarrow, Protter and Shimbo will be briefly recalled in section 3.2. But let us already point out that for type 3 bubbles, whether or not a bubble exists amounts to determining if the price process under a risk neutral measure is a martingale or a strict local martingale: if it is a strict local martingale, there is a bubble. The difference between a martingale and a strict local martingale has been recently investigated by several authors ([19], [112], [99] and [13] for instance). However, the distinction is subtle and in the case of a diffusion it amounts to understanding the asymptotic behavior of the asset price volatility. If the asset price volatility is large enough, then a bubble exists.

More formally, we model the asset price process by a standard stochastic differential equation driven by a Brownian motion W:

$$dS_t = \sigma(S_t)dW_t + \mu(S_t)dt \tag{3.1}$$

for all t in $[0, T]$, in some filtered probability space $(\Omega, \mathcal{F}, P, \mathbb{F})$ where $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$. We make the standing assumption that the asset price S is nonnegative. The asset price volatility $\sigma(S_t)$ is stochastic since it depends on the level of the asset price. Assuming no arbitrage in the sense of *No Free Lunch Vanishing Risk* (NFLVR), there exists a risk neutral measure (see [34]) under which this SDE simplifies to

$$S_t = S_0 + \int_0^t \sigma(S_s)dW_s. \tag{3.2}$$

It is well known (see Theorem 49) that this process S is a strict local martingale if and only if

$$\int_\alpha^\infty \frac{x}{\sigma^2(x)}dx < \infty \tag{3.3}$$

for all $\alpha > 0$. This last condition forms the basis of our bubble testing methodology, i.e. type 3 bubbles exist if and only if this integral is finite. The intuition behind the distinction between a martingale and a strict local martingale (in the case where the local martingale $S > 0$) derived from the fact that S is always a supermartingale and is a martingale if and only if it has constant expectation. So for a strict local martingale its expectation decreases with time. Thus on average under the risk neutral measure the buy and hold strategy is a losing one. A *typical* path of such a nonnegative continuous local martingale is to shoot up to high values and then quickly decrease to small values and remain at them; and this is also the typical behavior of prices of assets undergoing speculative bubbles.

We can also replace the equation (3.6) for S with a more general one, for example

$$\begin{aligned} dS_t &= \sigma(S_t)dW_t + \mu(S_t, Y_t)dt \\ dY_t &= s(Y_t)dB_t + g(Y_t)dt \end{aligned} \tag{3.4}$$

where B is a Brownian motion which is independent of W. This gives a model for S in the context of an incomplete market. This is perhaps the simplest model that implies an incomplete market. An alternative incomplete market model, and one that we do not consider here, introduces a stochastic volatility function as well. In any event, for our purposes, we are inevitably led in both situations (complete or incomplete models) to equation (3.7), under any risk neutral measure. If the models are complete, we appeal to the NFLVR framework of Cox and Hobson [31], and if the models are incomplete we can use the No Dominance framework of [83],[84], which we consider a better framework for models of bubbles.

Many authors have proposed estimators for the volatility function $\sigma(x)$. D. Florens-Zmirou [49] proposed a non parametric estimator based on the local time of the diffusion process. V. Genon Catalot and J. Jacod [52] propose an estimation procedure for parameterized volatility functions. M. Hoffmann [66] constructs a wavelets based estimator. In the second part of this chapter, we recall Florens-Zmirou's results. Since the constraint on the grid step, noted h_n in the sequel, required by Florens-Zmirou's theorem cannot be satisfied due to the limited data available, we propose another local time based estimator, using a smooth kernel, where the condition on h_n is easier to satisfy.

The main difficulty in using non parametric estimators is that one can estimate $\sigma(x)$ only at points visited by the process. We, therefore, cannot know the tails of the volatility function and determine if the integral in (3.3) is finite or infinite. In the third part of this chapter we propose two methods to deal with this *extrapolation problem*. The first method is based on a comparison theorem. We compare the behavior of parametric and non-parametric estimators of $\sigma(x)$. When the two estimators are statistically similar within the observation interval, we extrapolate into the tails using the parametric form's asymptotic behavior. The second method is based on Reproducing Kernel Hilbert Spaces (*RKHS*) theory. In fact, a roughly analogous problem arises in physical chemistry for potentials whose asymptotic behavior is known (see [67]). In our case, we do not know the asymptotic behavior of the volatility (that is what we are looking for!). To overcome this problem, we introduce a parameterized family (H_m) of RKHS's. Different m's allow us to construct interpolating functions with different asymptotic behaviors. We optimize over m in a sense that will be explained below and identify the reproducing kernel Hilbert space which allows us to construct an interpolating function that extends the non-parametric estimator from the observation interval to the entire real line.

We devote the last section to illustrating these various estimation methodologies. We focus on stocks from the alleged internet dotcom bubble of 1998 - 2001 (see for instance [125] and [110]) for which we could find relevant tick data. We selected four stocks: *Lastminute.com*, *Etoys*, *Infospace* and *Geocities*. The data was obtained from WRDS [126]. We use our methodology to see whether these stocks exhibited price bubbles. The evidence supports the existence of price bubbles. In addition, these four stocks allow us to illustrate the strengths and weaknesses of our testing methodology. During May 2011, there was speculation in the financial press concerning the existence of a price bubble in the aftermath of the recent IPO of LinkedIn. We analyzed stock price tick data from the short lifetime of this stock through May 24, 2011, and we found that LinkedIn has a price bubble.

An outline for this chapter is as follows. Section 3.2 presents the martingale theory of bubbles as developed by Jarrow, Protter and Shimbo. Section 3.3 is the first step in the bubbles detection methodology within a diffusion framework and presents the Florens-Zmirou's and smooth kernel volatility estimators on a compact domain. Section 3.4 extends these estimators to the nonnegative real line and presents our methodology for bubbles detection. Section 3.4.4 illustrates our testing methodology for asset price bubbles during the dotcom bubbles and 3.5.1 used our methodology to test in real time whether LinkedIn's stock price was exhibiting a bubble during May 2011. In Subsection 3.5.2, we attached two articles from the financial press that referenced during May and June 2011 our work on the real time detection of LinkedIn's bubble.

3.2 The martingale theory of asset bubbles

Due to the 2007 credit crisis, asset price bubbles have recently received considerable attention in the financial press, especially with respect to residential housing, commercial real estate, oil and gold prices. Concurrently, but initially independently, a new theory for understanding asset price bubbles has been growing in the academic literature. This approach, which was labeled the *martingale theory for bubbles*, defines a price bubble to exist when the asset price process is a strict local martingale (see Cox and Hobson [31], Jarrow, Protter and Shimbo [83],[84], Loewenstein and Willard [108]). We consider a continuous trading finite horizon economy $[0, T]$ with frictionless (no transaction costs) and competitive (all traders act as price takers believing that their trades do not impact the market price) markets. We focus on a finite time horizon where only type 3 bubbles can arise.

Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a filtered complete probability space and assume that $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}$ satisfies the usual hypothesis. We assume that a risky asset and a money market account are traded in our economy, and we assume for simplicity that the risk free interest rate is zero. See [82] for a complete exposition of the theory in the general case. Let $\tau \leq T$ be a stopping time which represents the maturity of the risky asset, and let $D = (D_t)_{0\leq t<\tau}$ be a nonnegative càdlàg semimartingale process adapted to $\mathbb{F}$ which represents the cumulative cash flow process from holding the risky asset. Let $X_\tau \in \mathcal{F}_\tau \geq 0$ be the time τ liquidation value of the asset. Assume that X_τ and $D_{\tau\wedge T}$ are integrable. Let the *market price* of the risky asset be given by the nonnegative càdlàg semimartingale $S = (S_t)_{0\leq t<\tau}$. Finally, let W be the wealth process associated with the market price of the risky asset plus its accumulated cash flows, that is

$$W_t = 1_{\{\tau>t\}} S_t + D_{t\wedge\tau} + X_\tau 1_{\{\tau\le t\}}$$

We assume in the following that there are no *arbitrage opportunities* in the sense of No Free Lunch with Vanishing Risk (NFLVR). By the first fundamental theorem of asset pricing (See [33] and [34]), we have the existence of an equivalent probability measure such that W is a Q local martingale. We will call any such Q an equivalent local martingale measure and we write ELLM. We assume that the market is incomplete and by the second fundamental theorem of asset pricing, such Q is not unique. To define the fundamental price of the asset, we need to choose a unique ELMM from the continuum of such possible measures. That this is possible is due to Jacod and Protter (see [75]) at least when enough call options are traded in the market. Now, given the market selected ELMM Q, we define the fundamental price of the risky asset as

$$S_t^* = E^Q\Big(\int_t^{\tau\wedge T} dD_u + X_\tau 1_{\{\tau\le T\}} \mid \mathcal{F}_t\Big)$$

This fundamental price represents the present value of the asset's future cash flows if held until liquidation. Under Q, $S_t^* + \int_0^t dD_u$ is a uniformly integrable martingale. We can now define the asset price bubble.

Definition 10 (Price Bubble) *The asset price bubble is the difference between the market price and the fundamental price: $S_t - S_t^*$. We write $\beta_t = S_t - S_t^*$ for the time t price bubble of S.*

The following holds. See [83] for a proof.

Theorem 47 *Any non-zero asset price bubble β is a Q strict local martingale such that*

(i) $\beta \ge 0$,

(ii) $\beta_\tau = 0$,

(iii) If $\beta_t = 0$, then $\beta_u = 0$ for each $u \ge t$, and

(iv) If there are no cash flows, then

$$S_t = E^Q(S_T \mid \mathcal{F}_t) + \beta_t - E^Q(\beta_T \mid \mathcal{F}_t)$$

In the sequel, we assume for simplicity that there are no cash flows and that $\tau = T$. In this case, for each $0 \le t \le T$,

$$\beta_t = E^Q(S_T \mid \mathcal{F}_t) - S_t$$

In order to value derivatives, one needs to introduce the No Dominance assumption, which was initially introduced by Merton. However, the theory for bubbles in contingent claims as developed in Jarrow, Protter and Shimbo, [83] uses implicitly that the market prices are local martingales under the probability measure Q. This is however not explicitly pointed out in any of the papers we referenced above. It is this assumption together with No Dominance that allowed the authors to prove that bounded payoffs have no bubbles and to derive consequently the call-put parity for the market prices. However, the assumption that all market prices are local martingales under the same measure Q appears to be very restrictive and constructing examples satisfying such assumption is very hard. So, contrarly to the work of Jarrow, Protter and Shimbo, we focus on the martingale theory for bubbles only for asset price bubbles and do not consider bubbles in derivatives, neither in theory nor in the practical bubbles detection.

Before focusing on the real time detection of asset bubble, we introduce in the next section some mathematical preliminaries. Will be mainly needed criteria to check the martingale property of a nonnegative local martingale, at least in a Brownian paradigm. The focus will be on stochastic and local volatility models. For local volatility models, a deterministic test to check the martingale property of the asset price is available and involves only the asymptotics of the volatility function. Therefore, in order to check in real time if type 3 asset bubbles within a local volatility framework exist, one needs to be at least able to estimate the volatility function. This is the aim of the second part of section 3.3, where non parametric volatility estimators are derived.

3.3 Mathematical preliminaries

3.3.1 Strict local martingales

The difference between a martingale and a strict local martingale has been recently investigated by several authors. We give now a general result that provides necessary and sufficient conditions for a nonnegative local martingale in a Brownian paradigm to be a true martingale. A partial list of relevant references is Sin [122], Kotani [99], Carr et al. [19] and Urusov and Mijatovic [112]. Usually, these results are based on explosions of solutions of SDEs and the martingale property of S is often equivalent to the non-explosion of some auxiliary process. For the rest of this subsection we assume that the filtration $\mathbb{F}$ is generated by a d-dimensional Brownian motion $W = (W^1, \ldots, W^d)$. We will need first the following very classical result. For completeness we provide a proof.

Lemma 77 *Let S be a strictly positive continuous local martingale. Then there is a Brownian motion B and a nonnegative predictable process ξ such that $\sqrt{\xi}$ is B-integrable and*

$$S_t = \mathcal{E}(X)_t, \quad \textit{where} \quad X_t = \int_0^t \sqrt{\xi_s} dB_s. \tag{3.5}$$

Proof. Since S is a strictly positive continuous local martingale, there exists a continuous local martingale X such that $S_t = \mathcal{E}(X)_t$. Using the assumption that $\mathbb{F}$ is a Brownian filtration, the representation theorem for continuous local martingales ensures the existence of a d-dimensional process $\eta \in L^{loc}(W)$ such that $X_t = \int_0^t \eta_s dW_s$. Hence $d\langle X \rangle_t = \xi_t dt$, where $\xi_t = ||\eta_t||^2 \geq 0$. Now define $B_t = \int_0^t \frac{1}{\sqrt{\xi_s}}\mathbf{1}_{\{\xi_s>0\}}dX_s + \int_0^t 1_{\{\xi_s=0\}}dW_s^1$. This process is well defined due to the finiteness of

$$E\left(\int_0^t \frac{1}{\xi_s}1_{\xi_s>0}d\langle X,X\rangle_s\right) = E(1_{\xi_s>0}ds) \leq t < \infty,$$

and is a continuous local martingale. Since it satisfies

$$\langle B,B\rangle_t = \int_0^t \frac{1}{\xi_s}\xi_s 1_{\xi_s>0}ds + \int_0^t 1_{\xi_s=0}ds = \int_0^t ds = t,$$

Lévy's theorem implies that B is in fact Brownian motion. One readily verifies that $dX_t = \sqrt{\xi}dB_t$, which proves the lemma. ■

We give now a characterization in the Brownian framework of the price processes S that are true martingales, under the assumption that ξ is non explosive under P. We know S is a supermartingale, hence for every $t \in [0,T]$, $E(S_t) \leq S_0$. Moreover, S is a martingale if and only if $E(S_t) = S_0$ for every t. Let n be an integer, define $\tau_n = \inf\{t \mid \xi_t \geq n\}$, and let $\tau = \lim_{n\to\infty}\tau_n$ be the explosion time of ξ. By assumption, ξ does not explode under P, hence $\tau = +\infty$ P-a.s. Clearly $(S_{t\wedge\tau_n})_{t\geq 0}$ is a strictly positive true martingale, and we can define measures Q^n by the Radon-Nikodym derivatives

$$\frac{dQ^n}{dP}\bigg|_{\mathcal{F}_{T\wedge\tau_n}} = \frac{S_{T\wedge\tau_n}}{S_0}.$$

This defines a consistent family of probability measures on $(\Omega, \mathcal{F}_{\tau_n\wedge T})$. To see this, let $n > m$ and let $B \in \mathcal{F}_{T\wedge\tau_m}$. Then $Q^n(B) = E^{Q^n}(1_B) = E^P(\frac{S_{T\wedge\tau_n}}{S_0}1_B) = E^P(E^P(\frac{S_{T\wedge\tau_n}}{S_0}1_B \mid \mathcal{F}_{T\wedge\tau_m})) = E^P(\frac{S_{T\wedge\tau_m}}{S_0}1_B) = Q^m(B)$, where the forth equality follows from the $\mathcal{F}_{T\wedge\tau_m}$ measurability of 1_B, the martingale property of $(S_{t\wedge\tau_n})_{t\geq 0}$ and the fact that the sequence of stopping times $(\tau_n)_{n\geq 1}$ is increasing. Therefore we can apply the Kolmogorov extension theorem (see for instance [69]), which ensures the existence and uniqueness of a measure Q^S on $(\Omega, \mathcal{F}_{\tau\wedge T}) = (\Omega, \mathcal{F}_T)$ such that $Q^S(B) = Q^n(B)$, $\forall\ B \in \mathcal{F}_{T\wedge\tau_n}$. The measure Q^S will be called the measure associated with the numeraire S. The following theorem is now classical, but we provide a proof for completeness.

Theorem 48 *Assume that ξ is non explosive under P. Then S is a P martingale if and only if ξ does not explode under Q^S, the measure associated with the numeraire S.*

Proof. Let $A \in \mathcal{F}_T$ and write

$$Q^S(A) = E^{Q^S}(1_A(1_{\{\tau\leq T\}} + 1_{\{\tau>T\}})) = Q^S(A\cap\{\tau \leq T\}) + E^{Q^S}(1_A 1_{\{\tau>T\}})$$

By the monotone convergence theorem,

$$E^{Q^S}(1_A 1_{\{\tau>T\}}) = E^{Q^S}(1_A \lim_{n\to\infty} 1_{\{\tau_n>T\}}) = \lim_{n\to\infty} E^{Q^S}(1_A 1_{\{\tau_n>T\}}).$$

Successive applications of the monotone convergence theorem and changes of measure give

$$\begin{aligned} E^{Q^S}(1_A 1_{\{\tau>T\}}) &= \lim_{n\to\infty} E^{Q^n}(1_A 1_{\{\tau_n>T\}}) = \lim_{n\to\infty} E^P(\frac{S_{T\wedge\tau_n}}{S_0} 1_A 1_{\{\tau_n>T\}}) \\ &= \lim_{n\to\infty} E^P(\frac{S_T}{S_0} 1_A 1_{\{\tau_n>T\}}) = E^P(\frac{S_T}{S_0} 1_A 1_{\{\tau>T\}}). \end{aligned}$$

We conclude that $Q^S(A) = Q^S(A\cap\{\tau\le T\}) + E^P(\frac{S_T}{S_0}1_A)$, and applying this with $A=\Omega$ we obtain that $1 = Q^S(\tau\le T) + E^P(\frac{S_T}{S_0})$. Thus $E^P(S_T) = S_0$ if and only if $Q^S(\tau\ge T) = 0$, and the criterion is proved. ■

This result is useful for stochastic volatility models, where the process ξ is specified through an SDE, and the explosion condition can be easily checked.

3.3.1.1 Stochastic volatility models

We consider the state space $J=(0,\infty)$ and a J-valued diffusion ν on some filtered probability space $(\Omega,\mathcal{F},\mathbb{F},P)$ governed by the SDE

$$d\nu_t = \sigma(\nu_t)dB_t + \mu(\nu_t)dt$$

Define S to be the solution to the SDE

$$dS_t = S_t b(\nu_t)dW_t$$

where W and Z are two correlated Brownian motions. We assume that the Engelbert-Schmidt conditions are satisfied ($\sigma(x)>0$ for all $x\in J$ and $\frac{1}{\sigma^2},\frac{\mu}{\sigma^2}\in L^1_{loc}(J)$), so that the SDE satisfied by ν has a unique in law (possibly explosive) weak solution. Let ξ be the explosion time of ν. Finally assume that $\frac{b^2}{\sigma^2}\in L^1_{loc}(J)$ so that $\int_0^t b^2(\nu_s)ds<\infty$ P a.s on $\{t<\xi\}$ so that the stochastic integral $\int_0^t b(\nu_s)dW_s$ is well defined for $t<\xi$. (See Mijatovic and Urusov, [112]). The following lemma holds as a consequence of the previous theorem.

Lemma 78 *If ν is not explosive, there exists an auxiliary J-valued process defined on some filtered probability space $(\Omega,\mathcal{G},\mathbb{G},Q)$ by the SDE*

$$dZ_t = \sigma(Z_t)dB_t + \big(\nu(Z_t) + \rho\sigma(Z_t)b(Z_t)\big)dt$$

such that $E^P(S_T) = S_0 Q(\tau>T)$ where τ is the explosion time of Z.

Remark 7 *The SDE that appears in the lemma above has a unique in law (possibly explosive) weak solution because the Engelbert-Schmidt conditions are again satisfied thanks to our assumption on b, μ and σ.*

Note that in [112], the authors provide a more general result, since they do not need that ν is not explosive. Their proofs are based on separating times techniques. We do present our weaker result only, since its proof is simple and in any case, this is not relevant for our asset bubbles detection problem, which is the primary aim of this chapter.

Define $v(x) = \int_c^x p'(y) \int_c^y \frac{2dz}{p'(z)\sigma^2(z)} dy$, where p is the scale function of Z. Using Feller's test of explosion, we obtain the following result from Lemma 78.

Lemma 79 *Assume ν is not explosive. Then S is a martingale if and only if $v(\infty) = v(0) = \infty$. In this case $E(S_t) = S_0$ and $\int_0^\infty E(S_t)dt = \infty$.*

Based on the values of v and p at zero and infinity, not only do we know if S is a martingale or a strict local martingale, but these same values give properties of the limit of $E(S_t)$ as t grows to infinity and can also give an idea about how fast $E(S_t)$ decays when S is a strict local martingale.

Lemma 80 *Assume that ν is not explosive and that $v(\infty) < \infty$ or $v(0) < \infty$. Then*

(i) S is a strict local martingale

(ii) $t \to E(S_t)$ is a non-increasing function

(iii) $\lim_{t\to\infty} E(S_t) = 0$ if and only if $P(\tau < \infty) = 1$.

So that if none of the following is satisfied (see [93, Proposition 5.32])

(a) $v(\infty) < \infty$ and $v(0) < \infty$

(b) $v(\infty) < \infty$ and $p(0) = -\infty$

(c) $v(0) < \infty$ and $p(\infty) = \infty$

then $\lim_{t\to} E(S_t) > 0$ (this is actually and equivalence) and $\int_0^\infty E(S_t)dt = \infty$.

Finally,

- *If (a) holds, then $\lim_{t\to\infty} E(S_t) = 0$ and $E(\tau) < \infty$ so that $\int_0^\infty E(S_t)dt = S_0 E(\tau) < \infty$.*
- *If (b) or (c) holds then $\lim_{t\to\infty} E(S_t) = 0$ but nothing can be concluded about $\int_0^\infty E(S_t)dt$.*

We turn now to local volatility models, for which similar results are available (see e.g. [19, 99, 122, 112].)

3.3.1.2 Local volatility models

The following theorem has been proved using different techniques. A proof based on explosion time techniques is provided in [19]. Kotani gives a PDE based proof in [99].

D. Kramkov has also recently pointed out how this follows simply from Feller's test for explosions [100].

Theorem 49 *If S solves the SDE*

$$dS_t = \sigma(S_t)dW_t$$

Then S is a true martingale (i.e. S has no price bubbles) if and only if $\int_{\varepsilon}^{\infty} \frac{x}{\sigma^2(x)} ds = \infty$, for all $\varepsilon > 0$.

The main step of the proof is to show that the martingale property of S is equivalent to the non-explosion to infinity of the solution to the auxiliary SDE

$$dZ_t = \frac{\sigma(Z_t)^2}{Z_t} dt + \sigma(Z_t)dB_t, \qquad Z_0 = S_0.$$

One then applies Feller's criterion for non-explosion to obtain the explicit condition that appears in Theorem 49. Note that the auxiliary process obtained above is of the same form as the one in the previous subsection, with $\mu = 0$, $b(x) = \frac{\sigma(x)}{x}$ and $\rho = 1$.

Example 8 *Consider the Constant Elasticity of Volatility (CEV) model*

$$dS_t = S_t^{\alpha} dW_t$$

where W is a standard Brownian motion. It follows from Theorem 49 that S is martingale if and only if $\alpha \leq 1$.

The next example is due to Sin, [122].

Example 9 *Let (Z, W) be a two-dimensional standard Brownian motion. Let (S, ν) satisfy*

$$\begin{aligned} dS_t &= S_t \nu_t^{\alpha} (\sigma_1 dZ_t + \sigma_2 dW_t) \\ d\nu_t &= \nu_t \big(\rho(b - \nu_t)dt + a_1 dZ_t + a_2 dW_t\big) \end{aligned}$$

under the ELMM Q. If $\alpha > 0$, $\rho \geq 0$, $b > 0$ and a_1, a_2, σ_1 and σ_2 are constants that satisfy $a_1\sigma_1 + a_2\sigma_2 > 0$, then S is a strict local martingale.

The condition in Theorem 49 forms the basis of our bubble testing methodology, i.e. type 3 bubbles exist if and only if this integral is finite. In order to check in real time if type 3 asset bubbles exist within a local volatility framework, we need to estimate first the volatility function. This is the aim of the next subsection where we propose non parametric volatility estimators.

3.3.2 Estimation of the volatility function in diffusion models

We model the asset price process by a standard stochastic differential equation driven by a Brownian motion $\overline{W}$:

$$dS_t = \sigma(S_t)d\overline{W}_t + \mu(S_t)dt \tag{3.6}$$

for all t in $[0, T]$, in some filtered probability space $(\Omega, \mathcal{F}, P, \mathbb{F})$ where $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$. We make the standing assumption that the asset price S is nonnegative. The asset price volatility $\sigma(S_t)$ is stochastic since it depends on the level of the asset price. Assuming no arbitrage in the sense of No Free Lunch Vanishing Risk (NFLVR), there exists a risk neutral measure (see [34]) under which this SDE simplifies to

$$S_t = S_0 + \int_0^t \sigma(S_s) dW_s. \tag{3.7}$$

where W is a Brownian motion under this risk neutral measure. We already pointed out that many authors have proposed estimators for the volatility function $\sigma(x)$. In the first part of this subsection, we recall Florens-Zmirou's results. Since the constraint on the grid step, noted h_n in the sequel, required by Florens-Zmirou's theorem cannot be satisfied due to the limited data available, we propose then another local time based estimator, using a smooth kernel, where the condition on h_n is easier to satisfy.

Note that the estimation is performed in the real (and not the risk neutral) world. However, the estimators we use do not involve the drift, hence without loss of generality, we assume throughout the remainder of the section that μ is identically null. Therefore, we consider the stochastic differential equation in (3.7) where the function $\sigma(x)$ is unknown. We define our non parametric estimator of $\sigma(x)$ based on discrete time observations $S_{t_1}, ..., S_{t_n}$, on the finite time interval $[0, T]$. We assume a regular sampling, that is $t_i = \frac{i}{n}T$.

3.3.2.1 Florens-Zmirou's Estimators

This section reviews Florens-Zmirou's estimators for our subsequent usage. Her estimator is based on the local time of a diffusion and is explained heuristically as follows. The local time is given by

$$\ell_T(x) = \lim_{\varepsilon\to 0} \frac{1}{2\varepsilon} \int_0^T 1_{\{|S_s - x| < \varepsilon\}} d\langle S, S\rangle_s$$

where $d\langle S, S\rangle_s = \sigma^2(S_s)ds$ so that $\ell_T(x) = \sigma^2(x)L_T(x)$, and

$$L_T(x) = \lim_{\varepsilon\to 0} \frac{1}{2\varepsilon} \int_0^T 1_{\{|S_s - x| < \varepsilon\}} ds.$$

Hence, the ratio $\frac{\ell_T(x)}{L_T(x)} = \sigma^2(x)$ yields the volatility at x. These limits and integrals can be approximated by the following sums :

$$L_T^n(x) = \frac{T}{2nh_n} \sum_{i=1}^n 1_{\{|S_{t_i} - x| < h_n\}}$$

$$\ell_T^n(x) = \frac{T}{2nh_n} \sum_{i=1}^n 1_{\{|S_{t_i} - x| < h_n\}} n(S_{t_{i+1}} - S_{t_i})^2$$

where h_n is a sequence of positive real numbers converging to 0 and satisfying some constraints. This allows us to construct an estimator of $\sigma(x)$ given by:

$$S_n(x) = \frac{\sum_{i=1}^n 1_{\{|S_{t_i}-x|<h_n\}} n(S_{t_{i+1}} - S_{t_i})^2}{\sum_{i=1}^n 1_{\{|S_{t_i}-x|<h_n\}}}. \tag{3.8}$$

Indeed, Florens-Zmirou [49] proves the following theorems.

Theorem 50 *If σ is bounded above and below from zero, has three continuous and bounded derivatives, and if $(h_n)_{n\geq 1}$ satisfies $nh_n \to \infty$ and $nh_n^4 \to 0$ then $S_n(x)$ is a consistent estimator of $\sigma^2(x)$.*

The proof of this theorem is based on the expansion of the transition density. The choice of a sequence h_n converging to 0 and satisfying $nh_n \to \infty$ and $nh_n^4 \to 0$ allows one to show that $L_T^n(x)$ and $\ell_T^n(x)$ converge in $L^2(dQ)$ to $L_T(x)$ and $\sigma^2(x)L_T(x)$, respectively. Hence $S_n(x)$ is a consistent estimator of $\sigma^2(x)$, for any x that has been visited by the diffusion.

We also have the following limit theorem, useful to obtain confidence intervals for the estimator $S_n(x)$ of $\sigma(x)$.

Theorem 51 *If moreover $nh_n^3 \to 0$ then $\sqrt{N_x^n}(\frac{S_n(x)}{\sigma^2(x)} - 1)$ converges in distribution to $\sqrt{2}Z$ where Z is a standard normal random variable and $N_x^n = \sum_{i=1}^n 1_{\{|S_{t_i}-x|<h_n\}}$.*

The aim of the next subsection is to construct another estimator based on the local time of the diffusion but using a smooth kernel. Theorem 50 will remain true under the constraint $nh_n^2 \to \infty$. This is important for the purpose of this chapter: requiring that h_n satisfies the conditions of Theorem 50 would provide useless estimators (not smooth enough to work with in practice) due to the limited data available to us. When testing our procedure, we will always provide these two estimators although theoretically we cannot be sure that the estimator of Florens-Zmirou converges due to the requirement that $h_n = n^{-\frac{1}{4}}$, which needs more data than we have.

3.3.2.2 A Smooth Kernel Estimator

We introduce a smooth kernel estimator to relax the condition on h_n to $nh_n^2 \to \infty$. In practice we often do not have enough data and the convergence condition $nh_n^4 \to 0$ is too restrictive. The key theorem of this section proves the convergence in probability of our sequence of smooth kernel estimators $S_n(x)$ to $\sigma^2(x)$.

For this section and without loss of generality, we assume that $T = 1$ and $t_i = \frac{i}{n}$. We consider again the discrete observation $S^{(n)} = (S_0, S_{\frac{1}{n}}, ..., S_1)$ defined through the stochastic differential equation (3.7). We assume that $\sigma(x)$ is bounded above and away from zero, C^3, and with bounded derivatives. These assumptions guarantee the existence of a unique strong solution which does not explode. We denote by Q the law on the space of continuous functions equipped with the canonical filtration $(\mathcal{F}_t)_{t\in[0,1]}$ and under which the canonical process $(S_t, 0 \leq t \leq 1)$ is a solution to the previous SDE.

Note that from a statistical point of view, it is more natural to work with weak solutions. The smoothness assumption and the boundedness of $\sigma(x)$ and its derivatives are required to obtain some estimates. We also consider a compact interval $\mathcal{D}$, which represents the observation interval, i.e. the domain on which the estimation is performed. We emphasize the fact that we are able to estimate $\sigma(x)$ only for those points that have been visited by the diffusion.

The idea underlying the smooth kernel estimator is to replace the kernel $K(x) = \frac{1}{2}1_{\{|x|\leq 1\}}$ by a smooth kernel ϕ, which is a C^6 positive function with compact support and such that $\int_{\mathbb{R}_+} \phi = 1$. We are interested in some L^p (p is stated later) convergence of the following quantities:

$$V_n^x = \frac{1}{nh_n}\sum_{i=0}^{n-1}\phi(\frac{S_{\frac{i}{n}} - x}{h_n})n(S_{\frac{i+1}{n}} - S_{\frac{i}{n}})^2 \tag{3.9}$$

$$L_n^x = \frac{1}{nh_n}\sum_{i=0}^{n-1}\phi(\frac{S_{\frac{i}{n}} - x}{h_n}) \tag{3.10}$$

to $\sigma^2(x)L^x$ and L^x respectively, where h_n satisfies $nh_n^2 \to \infty$.

Our convergence theorem requires the use of various lemmas. The first lemma involves the convergence of L_n^x and it follows from Proposition 3 of Hoffmann [66, p. 468].

Lemma 81 *Assume ϕ given in (3.10) is taken to be C^3. For each $\gamma \geq 2$, there exists a constant C such that*

$$\sup_{x\in\mathcal{D}} E(|L_n^x - L^x|^\gamma) \leq C(h_n^{\frac{\gamma}{2}} + (\frac{1}{nh_n^2})^\gamma).$$

Hence the L^p convergence of L_n^x to L^x is guaranteed, for all $p \geq 2$ and all $x \in \mathcal{D}$, and the L^p convergence for all $p > 0$ follows by Cauchy Schwarz. Our next lemma involves the L^1 convergence of V_n^x.

Lemma 82 *For each $x \in \mathcal{D}$, V_n^x converges in L^1 to $\sigma^2(x)L^x$.*

In order to prove this lemma, we write $V_n^x - \sigma^2(x)L^x = A_n(x) + B_n(x)$ where:

$$A_n(x) = \frac{1}{h_n}\int_0^1 \phi(\frac{X_s - x}{h_n})\sigma^2(X_s)ds - \sigma^2(x)L^x$$

$$B_n(x) = V_n^x - \frac{1}{h_n}\int_0^1 \phi(\frac{X_s - x}{h_n})\sigma^2(X_s)ds$$

We study each of those two terms separately. Let x be fixed in $\mathcal{D}$, where $\mathcal{D}$ is the domain over which the estimation is performed. Since x is fixed, we omit it from now on and prove that A_n and B_n converge in L^1 to 0.

Lemma 83 (Study of A_n) *For each $\gamma \geq 2$, there exists $c > 0$ such that $E|A_n|^\gamma \leq c h_n^{\frac{\gamma}{2}}$.*

Proof. Let $l = (l_1^x)$ be the local time of the diffusion at time $t = 1$. First, we use the occupation time formula,

$$\begin{aligned} A_n &= \frac{1}{h_n}\int_0^1 \phi(\frac{X_s - x}{h_n})\sigma^2(X_s)ds - \sigma^2(x)L^x \\ &= \frac{1}{h_n}\int_{\mathbb{R}+} \phi(\frac{y-x}{h_n})l^y dy - \frac{1}{h_n}\int_{\mathbb{R}+} \phi(\frac{y-x}{h_n})l^x dy = \frac{1}{h_n}\int_{\mathbb{R}+} \phi(\frac{y-x}{h_n})(l^y - l^x)dy. \end{aligned}$$

Applying Jensen's lemma to the integral, a straightforward change of variables and taking expectations give:

$$E|A_n|^\gamma \leq \frac{1}{h_n^\gamma} E(\int_{\mathbb{R}+} |l^{zh_n+x} - l^x|^\gamma \phi^\gamma(z) h_n^\gamma dz).$$

The following inequalities follow from an application of Fubini's theorem and the Hölder property of the local time paths of a continuous local martingale. (This is a well known and classic result, given for example by Revuz and Yor [119], p. 227.)

$$E|A_n|^\gamma \leq \int_{\mathbb{R}+} \phi^\gamma(z) E(|l^{zh_n+x} - l^x|^\gamma)dz \leq h_n^{\frac{\gamma}{2}} \int_{\mathbb{R}+} |z|^{\frac{\gamma}{2}} \phi^\gamma(z)dz$$

Since ϕ has compact support, we can take $c = \int_{\mathbb{R}+} |z|^{\frac{\gamma}{2}}\phi^\gamma(z)dz$. Then $E|A_n|^\gamma \leq c h_n^{\frac{\gamma}{2}}$ and the lemma is proved. ■

We now focus on the study of B_n, which we write $-B_n = C_n + D_n$ where:

$$\begin{aligned} C_n &= \frac{1}{h_n}\sum_{i=0}^{n-1}\int_{\frac{i}{n}}^{\frac{i+1}{n}} (\phi(\frac{S_s - x}{h_n})\sigma^2(S_s) - \phi(\frac{S_{\frac{i}{n}} - x}{h_n})\sigma^2(S_{\frac{i}{n}}))ds \\ D_n &= \frac{1}{h_n}\sum_{i=0}^{n-1}\int_{\frac{i}{n}}^{\frac{i+1}{n}} \phi(\frac{S_{\frac{i}{n}} - x}{h_n})(\sigma^2(S_{\frac{i}{n}}) - n(S_{\frac{i+1}{n}} - S_{\frac{i}{n}})^2)ds \end{aligned}$$

We need the following lemma borrowed from Genon-Catalot and Jacod, which we recall in the one dimensional setting. We refer to [52] for a proof. We define $X_i^n = \sqrt{n}\sigma(S_{\frac{i}{n}})(W_{\frac{i+1}{n}} - W_{\frac{i}{n}})$ and $Y_i^n = \sqrt{n}\int_{\frac{i}{n}}^{\frac{i+1}{n}} \sigma(S_s)dW_s$.

Lemma 84 (Genon-Catalot and Jacod) *Let $g \in C^2$. Assume there exists $\gamma > 0$ such that for all x, $|g(x)| + |g'(x)| + |g''(x)| \leq \gamma(1 + |x|^\gamma)$. Then there exists a constant C such that*

$$E\Big(\big(g(X_i^n) - g(Y_i^n)\big)^2 |\mathcal{F}_{\frac{i}{n}}\Big) \leq \frac{C}{n}.$$

If g is an even function, $\big|E(g(X_i^n) - g(Y_i^n)|\mathcal{F}_{\frac{i}{n}})\big| \leq \frac{C}{n}$.

We can now study the convergence of D_n.

Lemma 85 (Study of D_n) *There exists $C > 0$, such that $E|D_n| \leq \frac{C}{n}E\Big(\frac{1}{nh_n}\sum_{i=0}^{n-1}\phi(\frac{S_{\frac{i}{n}}-x}{h_n})\Big)$ and D_n converges to 0 in L^1.*

Proof. Recall that $D_n = \frac{1}{h_n}\sum_{i=0}^{n-1}\phi(\frac{S_{\frac{i}{n}}-x}{h_n})\big(\frac{\sigma^2(S_{\frac{i}{n}})}{n} - \big(\int_{\frac{i}{n}}^{\frac{i+1}{n}}\sigma(S_s)dW_s\big)^2\big)$. Write $g(x) = x^2$. The following inequalities are straightforward.

$$\begin{aligned}
E|D_n| &\leq \frac{1}{h_n}\sum_{i=0}^{n-1}E\Big(\phi(\frac{S_{\frac{i}{n}}-x}{h_n})\big|E\big(\frac{\sigma^2(S_{\frac{i}{n}})}{n}|\mathcal{F}_{\frac{i}{n}}\big) - E\big((\int_{\frac{i}{n}}^{\frac{i+1}{n}}\sigma(S_s)dW_s)^2|\mathcal{F}_{\frac{i}{n}}\big)\big|\Big)\\
&\leq \frac{1}{h_n}\sum_{i=0}^{n-1}E\Big(\phi(\frac{S_{\frac{i}{n}}-x}{h_n})\big|E(g(X_i^n)-g(Y_i^n)|\mathcal{F}_{\frac{i}{n}})\big|\Big)\frac{1}{n}.
\end{aligned}$$

Since g is an even function, Lemma 84 ensures the existence of a constant C such that:

$$\big|E\big(g(X_i^n)-g(Y_i^n)|\mathcal{F}_{\frac{i}{n}}\big)\big| \leq \frac{C}{n}.$$

Hence, $E|D_n| \leq \frac{C}{n}E\Big(\frac{1}{nh_n}\sum_{i=0}^{n-1}\phi(\frac{S_{\frac{i}{n}}-x}{h_n})\Big)$. Using Lemma 81, the sum converges to the local time of the diffusion in x and $E|D_n| \to 0$. ■

In order to study C_n, we introduce the function $f(y) = \phi(\frac{y-x}{h_n})\sigma^2(y)$ which is C^3 by assumption. We use a third order Taylor expansion and get: for all s in $[\frac{i}{n}, \frac{i+1}{n}[$, there exists $\Xi_{s,\frac{i}{n}}$ such that

$$f(S_s) = f(S_{\frac{i}{n}}) + (S_s - S_{\frac{i}{n}})f^{'}(S_{\frac{i}{n}}) + \frac{(S_s-S_{\frac{i}{n}})^2}{2}f^{''}(S_{\frac{i}{n}}) + \frac{(S_s-S_{\frac{i}{n}})^3}{6}f^{(3)}(\Xi_{s,\frac{i}{n}}).$$

We plug this into C_n and obtain $C_n = C_n^1 + C_n^2 + C_n^3 + C_n^4$ where

$$\begin{aligned}
C_n^1 = \frac{1}{h_n}\sum_{i=0}^{n-1}\int_{\frac{i}{n}}^{\frac{i+1}{n}}\Big(&(S_s-S_{\frac{i}{n}})\phi(\frac{S_{\frac{i}{n}}-x}{h_n})(\sigma^2)^{'}(S_{\frac{i}{n}}) + \frac{1}{2}(S_s-S_{\frac{i}{n}})^2\Big(\frac{2}{h_n}\phi^{'}(\frac{S_{\frac{i}{n}}-x}{h_n})\\
&(\sigma^2)^{'}(S_{\frac{i}{n}}) + \phi(\frac{S_{\frac{i}{n}}-x}{h_n})(\sigma^2)^{''}(S_{\frac{i}{n}})\Big) + \frac{1}{6}(S_s-S_{\frac{i}{n}})^3\Big(\frac{3}{h_n^2}\phi^{''}(\frac{\Xi_{s,\frac{i}{n}}-x}{h_n})(\sigma^2)^{'}(\Xi_{s,\frac{i}{n}})\\
&+\frac{3}{h_n}\phi^{'}(\frac{\Xi_{s,\frac{i}{n}}-x}{h_n})(\sigma^2)^{''}(\Xi_{s,\frac{i}{n}}) + \phi(\frac{\Xi_{s,\frac{i}{n}}-x}{h_n})(\sigma^2)^{(3)}(\Xi_{s,\frac{i}{n}})\Big)\Big)ds
\end{aligned}$$

and

$$\begin{aligned}
C_n^2 &= \frac{1}{h_n^2}\sum_{i=0}^{n-1}\phi^{'}(\frac{S_{\frac{i}{n}}-x}{h_n})\sigma^2(S_{\frac{i}{n}})\int_{\frac{i}{n}}^{\frac{i+1}{n}}(S_s-S_{\frac{i}{n}})ds\\
C_n^3 &= \frac{1}{h_n^3}\sum_{i=0}^{n-1}\phi^{''}(\frac{S_{\frac{i}{n}}-x}{h_n})\sigma^2(S_{\frac{i}{n}})\int_{\frac{i}{n}}^{\frac{i+1}{n}}\frac{(S_s-S_{\frac{i}{n}})^2}{2}ds\\
C_n^4 &= \frac{1}{h_n^4}\sum_{i=0}^{n-1}\int_{\frac{i}{n}}^{\frac{i+1}{n}}\phi^{(3)}(\frac{\Xi_{s,\frac{i}{n}}-x}{h_n})\sigma^2(\Xi_{s,\frac{i}{n}})\frac{(S_s-S_{\frac{i}{n}})^3}{6}ds
\end{aligned}$$

We prove in the following lemma that C_n^1 converges in L^1 to 0 when nh_n^2 tends to infinity. The idea is to bound C_n^1 by first bounding σ, ϕ and their three first derivatives and then using the Burkholder-Davis-Gundy inequalities (hereafter referred to simply as BDG) to obtain estimates of powers of $S_s - S_{\frac{i}{n}}$.

Lemma 86 (Study of C^1) *Assume that $nh_n^2 \to \infty$. Then C_n^1 converges in L^1 to 0.*

Proof. Since ϕ and σ and their derivatives are bounded, there exist nonnegative constants $(c_i)_{1\le i\le 6}$ such that

$$E|C_n^1| \le E\Big(\sum_{i=0}^{n-1}\frac{c_1}{h_n}\int_{\frac{i}{n}}^{\frac{i+1}{n}}|S_s - S_{\frac{i}{n}}|ds + (\frac{c_2}{h_n}+\frac{c_3}{h_n^2})\int_{\frac{i}{n}}^{\frac{i+1}{n}}|S_s - S_{\frac{i}{n}}|^2 ds$$
$$+(\frac{c_4}{h_n}+\frac{c_5}{h_n^2}+\frac{c_6}{h_n^3})\int_{\frac{i}{n}}^{\frac{i+1}{n}}|S_s - S_{\frac{i}{n}}|^3 ds\Big)$$

It follows clearly that

$$E|C_n^1| \le \sum_{i=0}^{n-1}\frac{c_1}{h_n}\int_{\frac{i}{n}}^{\frac{i+1}{n}}E(\sup_{s\in[\frac{i}{n},\frac{i+1}{n}]}|S_s - S_{\frac{i}{n}}|)ds + (\frac{c_2}{h_n}+\frac{c_3}{h_n^2})\int_{\frac{i}{n}}^{\frac{i+1}{n}}E(\sup_{s\in[\frac{i}{n},\frac{i+1}{n}]}|S_s - S_{\frac{i}{n}}|^2)ds$$
$$+(\frac{c_4}{h_n}+\frac{c_5}{h_n^2}+\frac{c_6}{h_n^3})\int_{\frac{i}{n}}^{\frac{i+1}{n}}E(\sup_{s\in[\frac{i}{n},\frac{i+1}{n}]}|S_s - S_{\frac{i}{n}}|^3)ds$$

We apply now BDG inequalities for continuous local martingales. For each $1 \le p \le 3$, there exist nonnegative constants C_p such that

$$E\Big(\sup_{s\in[\frac{i}{n},\frac{i+1}{n}]}|S_s - S_{\frac{i}{n}}|^p\Big) \le E\Big(\big(\int_{\frac{i}{n}}^{\frac{i+1}{n}}\sigma^2(S_s)ds\big)^{\frac{p}{2}}\Big) \le \frac{C_p}{n^{\frac{p}{2}}}.$$

Integrating, summing, and taking the expectation, we finally obtain

$$E|C_n^1| \le \frac{c_1}{h_n}\frac{C_1}{\sqrt{n}} + (\frac{c_2}{h_n}+\frac{c_3}{h_n^2})\frac{C_2}{n} + (\frac{c_4}{h_n}+\frac{c_5}{h_n^2}+\frac{c_6}{h_n^3})\frac{C_3}{n\sqrt{n}}$$
$$\le \frac{c_1C_1}{\sqrt{nh_n^2}} + \frac{C_2}{nh_n^2}(c_2h_n + c_3) + \frac{C_3}{nh_n^2}(\frac{c_4h_n}{\sqrt{n}} + \frac{c_5}{\sqrt{n}} + \frac{c_6}{\sqrt{nh_n^2}})$$

Since $nh_n^2 \to \infty$ and $h_n \to 0$, it follows clearly that C_n^1 converges in L^1 to 0. ■

We turn now to the study of C_n^2.

Lemma 87 (Study of C_n^2) *If $nh_n^2 \to \infty$, then C_n^2 converges in L^1 to 0.*

Proof. First, $E|C_n^2| \le \frac{1}{h_n^2}||\sigma^2||_\infty E\Big(\sum_{i=0}^{n-1}|\phi'|(\frac{S_{\frac{i}{n}}-x}{h_n})\Big|E\Big(\int_{\frac{i}{n}}^{\frac{i+1}{n}}|S_s - S_{\frac{i}{n}}|ds|\mathcal{F}_{\frac{i}{n}}\Big)\Big|\Big)$. It follows from an application of a BDG inequality that there exists a constant M_1 such that

$$E|C_n^2| \le \frac{M_1}{\sqrt{n}h_n}E\Big(\frac{1}{nh_n}\sum_{i=0}^{n-1}|\phi'(\frac{S_{\frac{i}{n}}-x}{h_n})|\Big).$$

Now using the kernel $\frac{1}{a}|\phi'|$, where $a = \int |\phi'|(x)dx$, the quantity $\frac{1}{nh_n}\sum_{i=0}^{n-1}|\phi'|(\frac{S_{\frac{i}{n}}-x}{h_n})$ converges in L^1 to aL^x. Hence $E|C_n^2|$ converges to zero as soon as $nh_n^2 \to \infty$, which proves the lemma. ■

We can provide another proof of this result following the proof in [66], page 477. We obtain that for each $\gamma \geq 2$ there exists a constant C such that

$$E|C_n^2|^\gamma \leq C\frac{1}{n^\gamma}h_n^{-\frac{3}{2}}$$

It follows that $E|C_n^2| \leq \sqrt{E(|C_n^2|^2)} \leq \frac{\sqrt{C}}{nh_n^2}h_n^{\frac{5}{4}}$ which converges to zero. The remaining terms to estimate are C_n^3 and C_n^4. Note that in the expansion of C_n, C_n^3 is the most important term.

Lemma 88 (Study of C_n^3 and C_n^4) *If $nh_n^2 \to \infty$, then C_n^3 and C_n^4 converge to zero in L^1.*

Proof. It is straigthforward to obtain the estimate

$$E|C_3^n| \leq \frac{M}{h_n^3}E\Big(\sum_{i=0}^{n-1}|\phi''|(\frac{S_{\frac{i}{n}}-x}{h_n})\int_{\frac{i}{n}}^{\frac{i+1}{n}}\frac{1}{2}E\big((S_s-S_{\frac{i}{n}})^2 \mid \mathcal{F}_{\frac{i}{n}}\big)\Big)$$

Now a BDG inequality implies the existence of a constant C such that $E((S_s - S_{\frac{i}{n}})^2 \mid \mathcal{F}_{\frac{i}{n}}) \leq C(s-\frac{i}{n})$. Hence

$$E|C_3^n| \leq \frac{MC}{2nh_n^2}E\big(\frac{1}{nh_n}\sum_{i=0}^{n-1}|\phi''|(\frac{S_{\frac{i}{n}}-x}{h_n})\big)$$

which converges to zero when $nh_n^2 \to \infty$. The same techniques as in this lemma and the previous one can be applied to prove the convergence of C_n^4 in L^1 to 0. ■

We have a stronger result that the one stated in the lemma above. Under the assumption that $nh_n^2 \to \infty$, both C_n^3 and C_n^4 converge to zero in L^γ, for all $\gamma > 2$. Define

$$\tilde{C}_n^3 = \frac{1}{h_n^3}\sum_{i=0}^{n-1}\phi''(\frac{S_{\frac{i}{n}}-x}{h_n})\int_{\frac{i}{n}}^{\frac{i+1}{n}}\frac{(S_s-S_{\frac{i}{n}})^2}{2}ds$$

Hoffman's result in [66] guarantees that $\tilde{C}_n^3$ converges in L^γ to 0 as soon as $nh_n^2 \to \infty$, for all $\gamma > 2$. Since σ and its derivatives are assumed to be bounded, it is not hard to see that we have the same result for C_n^3.

Proof. We follow again the steps in [66]. Define

$$\tilde{C}_n^3 = \frac{1}{h_n^3}\sum_{i=0}^{n-1}\phi''(\frac{S_{\frac{i}{n}}-x}{h_n})\int_{\frac{i}{n}}^{\frac{i+1}{n}}\frac{(S_s-S_{\frac{i}{n}})^2}{2}ds$$

Hoffman's result guarantees that $\tilde{C}_n^3$ converges in L^γ to 0 as soon as $nh_n^2 \to \infty$, for all $\gamma > 2$. Since σ and its derivatives are assumed to be bounded, it is not hard to see that we have the same result for C_n^3. Using Itô's formula, we have:

$$(S_s - S_{\frac{i}{n}})^2 = \int_{\frac{i}{n}}^{s} (S_u - S_{\frac{i}{n}})\sigma(S_u)dW_u + \int_{\frac{i}{n}}^{s} \sigma^2(S_u)du$$

We can then write : $C_n^3 = \frac{1}{h_n^3}\sum_{i=0}^{n-1}(T_{i,1} + T_{i,2})$ where

$$T_{i,1} = \frac{1}{2}\phi''(\frac{S_{\frac{i}{n}} - x}{h_n})\sigma(S_{\frac{i}{n}})^2 \int_{\frac{i}{n}}^{\frac{i+1}{n}} ds \int_{\frac{i}{n}}^{s} (S_u - S_{\frac{i}{n}})\sigma(S_u)dW_u$$
$$T_{i,2} = \frac{1}{2}\phi''(\frac{S_{\frac{i}{n}} - x}{h_n})\sigma(S_{\frac{i}{n}})^2 \int_{\frac{i}{n}}^{\frac{i+1}{n}} (\frac{i+1}{n} - s)\sigma^2(S_u)du$$

The second term is actually easy to bound : $|T_{i,2}| \leq M\big|\phi''(\frac{S_{\frac{i}{n}}-x}{h_n})\big|\frac{1}{n^2}$ and for $\gamma > 2$,

$$\frac{1}{h_n^{3\gamma}}E\Big(\big|\sum_{i=0}^{n-1} T_{i,2}\big|^\gamma\Big) \leq \frac{M}{h_n^{3\gamma}}\frac{1}{n^\gamma}E\Big(\Big|\sum_{i=0}^{n-1}\big|\phi''(\frac{S_{\frac{i}{n}} - x}{h_n})\big|\frac{1}{n}\Big|^\gamma\Big) \leq \frac{M}{(nh_n^2)^\gamma}$$

To deal with the first term, we define: $\tilde{T}_{i,1} = \frac{1}{2}\phi''(\frac{S_{\frac{i}{n}}-x}{h_n})\int_{\frac{i}{n}}^{\frac{i+1}{n}} ds \int_{\frac{i}{n}}^{s}(S_u - S_{\frac{i}{n}})\sigma(S_u)dW_u$. Using the Burkholder-Rosenthal inequality (see, e.g., [64]),

$$E\Big(\big|\sum_{i=0}^{n-1} T_{i,1}\big|^\gamma\Big) \leq C\bigg(\Big(E(\sum_{i=0}^{n-1} E(T_{i,1}^2|\mathcal{F}_i^n))\Big)^{\frac{\gamma}{2}} + \sum_{i=0}^{n} E|T_{i,1}|^\gamma\bigg).$$

Next by the boundedness of σ and its derivatives, we get:

$$E\Big(\big|\sum_{i=0}^{n-1} T_{i,1}\big|^\gamma\Big) \leq M\bigg(\Big(E(\sum_{i=0}^{n-1} E(\tilde{T}_{i,1}^2|\mathcal{F}_i^n))\Big)^{\frac{\gamma}{2}} + \sum_{i=0}^{n} E\big|\tilde{T}_{i,1}\big|^\gamma\bigg)$$

The right side quantity can be bounded by $Ch_n^\gamma\frac{1}{n^\gamma}$ (see [66] for a proof). Hence, putting all this together, we obtain that $E|C_n^3|^\gamma \leq \frac{M}{(nh_n^2)^\gamma}$. The same techniques as in this lemma and the previous one can be adapted to prove the convergence of C_n^4 in L^γ to 0 for all $\gamma > 2$. ■

Putting all these lemmas together proves that C_n converges in L^1 to 0, and thus B_n converges in L^1 to 0. We have then proved that L_n^x converges to L^x in L^p, for all $p > 0$ and that $V_n^x - \sigma^2(x)L^x$ converges in L^1 to 0, which ends the proof of Lemma 82. The following theorem is now straightforward.

Theorem 52 *If $nh_n^2 \to \infty$ then $S_n^x = \frac{V_n^x}{L_n^x}$ converges in probability to $\sigma^2(x)$ and provides a consistent estimator of $\sigma^2(x)$.*

Remark 8 *After finding and proving this theorem, we learned to our chagrin, from Jean Jacod, that he had not only already considered this exact problem more than 10 years ago, but that he has also established similar results which are both more general and more effective. See [73] and [74]. In particular he is able to take* $h_n = \frac{1}{\sqrt{n}}$ *and he also obtains a rate of convergence and an associated Central Limit Theorem. We have decided nevertheless to retain our estimator and its proof presented here, since it is the one we used to process the data and it seems to works well for our purposes. But we wish to signal for future related work that there are more powerful (if perhaps slightly more complicated) similar estimators available. These remarks also apply for parts of Section 3.3.2.3.*

3.3.2.3 Unbounded Volatility Function Estimators

The previous two estimators for the volatility function $\sigma(x)$ are over a compact domain representing the observation interval. In this section, for the SDE (3.7), we relax this boundedness assumption on the volatility function $\sigma(x)$. Herein, we now assume that $\sigma > 0$ on $I =]0, \infty[$, it is identically null elsewhere and satisfies $\frac{1}{\sigma^2} \in L^1_{loc}(I)$.

This is the Engelbert Schmidt condition (see, e.g., [46] or [93]) under which the SDE has a unique weak solution S that does not explode to ∞. We let P be the law of the solution on the canonical space $\Omega = C([0,T], \mathbb{R})$ equipped with the canonical filtration $(\mathcal{F}_t)_{t\in[0,T]}$ and the canonical process $S = (S_t)_{t\in[0,T]}$. We also assume that σ is C^3, and therefore it is bounded and with bounded derivatives on every compact set. We add in passing that these hypotheses imply the existence of a strong solution, as well. Let $\tau_0(S)$ be the first time S hits zero. The following theorem provides straightforward but useful extensions of Theorem 50 and Theorem 52.

Theorem 53 *Suppose* $\sigma(x)$ *has three continuous derivatives. Assume that* $nh_n^4 \to 0$ *and* $nh_n \to \infty$*. Then conditional on* $\{\tau_0(S) > T\}$*,* $S_n(x)$ *given in (3.8) converges in probability to* $\sigma^2(x)$*. The same holds for our smooth kernel estimator under the constraint* $nh_n^2 \to \infty$*.*

Proof. Let $T_q = \inf\{t, S_t \geq q\}$ and $\tau_p = \inf\left\{t, S_t \leq \frac{1}{p}\right\}$. Then $\lim_{p\to\infty} \tau_p = \tau_0(S)$ and $\lim_{q\to\infty} T_q = \infty$ since S does not explode to ∞. We can take $\sigma_{p,q}$ to be a function bounded above and below away from zero with three bounded derivatives such that $\sigma_{p,q}(x) = \sigma(x)$ for all $\frac{1}{p} \leq x \leq q$. Let $(S_t^{p,q})_{t\in[0,T]}$ be the unique strong solution to the SDE $dS_t^{p,q} = \sigma_{p,q}(S_t^{p,q})dW_t$. Introduce now $S_n^{p,q}(x)$, the estimator computed on the basis of $(S_t^{p,q})_{t\in[0,T]}$ as in (3.8) or using our smooth kernel estimator. Then under the suitable constraints on the sequence $(h_n)_{n\geq 1}$, $S_n^{p,q}(x)$ converges in probability to $\sigma^2_{p,q}(x)$. Moreover $S_n^{p,q}(x) = S_n(x)$ if $T < T_q \wedge \tau_p$. Then obviously $S_n(x)$ converges in probability to $\sigma^2(x)$, in restriction to the set $\{T < \tau_0(S)\}$. ∎

We can extend Theorem 51 by working in the filtration $\mathcal{G}_t = \sigma(S_s, s \leq t) \vee \sigma(\tau_p)$ where $\tau_p = inf\left\{t, S_t \in]\frac{1}{p}, p[\right\}$ and more interestingly in $\mathcal{G}'_t = \sigma(S_s, s \leq t) \vee U$ where $U = 1_{\{\tau_p > T\}}$ and T is the time horizon. Consider for instance the initial enlargement $\mathcal{G}_t = \mathcal{F}_t \vee \sigma(\tau_p)$ where $\mathcal{F}_t = \sigma(S_s, s \leq t)$ and τ_p is the first exit time of S from $]\frac{1}{p}, p[$. τ_p is an $\mathcal{F}$ stopping time. Consider the filtered probability space $(\Omega, \mathcal{G}, Q, \mathbb{G})$ where $\mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$. Under some technical assumptions, and if the sequence h_n satisfies $h_n \to 0$, $nh_n \to \infty$ and $nh_n^3 \to 0$ then we can prove that $1_{\{\tau_p > t\}}\sqrt{N_a^n}(\frac{S_n(a)}{\sigma^2(a)} - 1)$ converges in distribution to $\sqrt{2}Z$ where Z is a standard normal random variable and $N_a^n = \sum_{i=1}^n 1_{\{|S_{\frac{i}{n}} - a| < h_n\}}$.

Remark 9 (In practice) *Note that this limit theorem can be applied if during the time interval* $[0, T]$ *the process does not hit 0. The limit theorem also provides us with a confidence interval for the volatility estimator.*

3.4 How to detect an asset bubble in real time

We illustrate our methodology using data from the alleged internet dotcom bubble of 1998-2001. Not surprisingly, we find that all three eventualities occur: in one case we are able to confirm the presence of a bubble; in a second case we confirm the lack of a bubble, and in a third case we find that the test is inconclusive. It is our hope that our methodology opens some new avenues for the testing of stock price bubbles in real time.

3.4.1 The methodology

Theorem 49 forms the basis for our bubbles testing methodology. Unfortunately, we are immediately faced with an extrapolation problem. To see this problem, we note that the volatility function estimators presented in the previous sections provide estimates for $\sigma(x)$ only on a finite interval - those x that have been visited by the process. Given the available stock price data, we can not observe the tails of the volatility function necessary to check for the divergence of the integral in Theorem 49. To check for divergence, we must extrapolate from the observed domain of $\sigma(x)$ to the entire nonnegative real line.

We propose two extrapolation methods to overcome this problem. The first method is to use a parametric estimator as in [52], and a comparison theorem to conclude when the parametric and non parametric volatility estimators are similar. If similar, we extrapolate into the tail using the parametric form's asymptotic behavior. The second method is to use Reproducing Kernel Hilbert Spaces theory to extrapolate the volatility function in the *best* possible way. We will quantify what we mean by 'best' below.

3.4.2 Method 1: Parametric Estimation

The appeal of using a parametric form for the volatility function $\sigma(x)$ is that we know the tails once the parameters have been estimated. For this estimation we choose a class of volatility functions large enough to include many of the forms used in practice (for example : power functions $\sigma(x) = \sigma x^\alpha$, where σ and α are the unknown parameters that we estimate). For this example of $\sigma(x) = x^\alpha$ we have the process S is a strict local martingale that is always strictly positive if $\alpha > 1$, and is a martingale if $\frac{1}{2} \leq \alpha < 1$ which however can assume the value 0. If $\alpha = 1$ we are in the case of geometric Brownian motion. We then also use our non-parametric estimators. If these estimators are comparable, we have a conclusive test for divergence of the volatility integral. If they are not comparable, then the test is inconclusive.

3.4.2.1 The Comparison Theorem

This section states and proves the comparison theorem.

Theorem 54 (Comparison Theorem) *Assume that $dS_t = \sigma(t, S_t)dW_t$ and that there exist two functions Σ and $\overline{\sigma}$ such that : for all t and x, $\overline{\sigma}(x) \leq \sigma(t,x) \leq \Sigma(x)$ and such that σ, Σ and $\overline{\sigma}$ are continuous, locally Hölder continuous with exponent $\frac{1}{2}$, then:*

(i) if for all $c > 0$, $\int_c^\infty \frac{x}{\Sigma^2(x)}dx = \infty$ then S is a martingale.

(ii) if there exists $c > 0$ such that $\int_c^\infty \frac{x}{\overline{\sigma}^2(x)}dx < \infty$ then S is a strict local martingale.

In order to prove this theorem, we need the following lemma (see [43]):

Lemma 89 *Let g be a concave function, α_i, $i = 1, 2$ be two continuous functions, locally Hölder continuous with exponent $\frac{1}{2}$ such that for all (x,t), $\alpha_1(x,t) \leq \alpha_2(x,t)$. Let $T > 0$ be fixed. We consider $dX_t^{\alpha_{1,2}} = \alpha_{1,2}(t, X_t^{\alpha_{1,2}})dW_t$ and $u_{1,2}(x,t) = E_{(x,t)}(g(X_T^{\alpha_{1,2}}))$. Then for all $x \in \mathbb{R}^+$ and $t \in [0,T]$, $u_1(x,t) \geq u_2(x,t)$.*

Proof. [of Theorem 54] (i) Since g(x)=x is concave, we can apply the previous lemma and get that for all (x,t), $u(x,t) \geq u^\Sigma(x,t)$. If $\int_c^\infty \frac{x}{\Sigma^2(x)}dx = \infty$, then by Theorem 49, $u^\Sigma(x,t) = x$, for all (x,t). Thus for all (x,t), $u(x,t) \geq x$. But, we know that S is positive local martingale and hence a super martingale by Jensen's lemma, thus for all (x,t), $u(x,t) \leq x$. This proves that : $E(S_T|S_t = x) = x$ and S is a martingale.

(ii) Let $T > 0$ be fixed, and $\overline{u}(x,t) = E(S_T^{\overline{\sigma}}|S_t^{\overline{\sigma}} = x)$. Let $c > 0$ such that $\int_c^\infty \frac{x}{\overline{\sigma}^2(x)}ds < \infty$. We know that $S^{\overline{\sigma}}$ is strict local martingale by Theorem 49 and $\overline{u}(x,t) \leq x$ and it exists t such that $\overline{u}(x,t) < x$. Again since $g(x) = x$ is concave, $x \geq \overline{u}(x,t) \geq u(x,t)$, for all (t,x), and there exists t such that $u(x,t) < \overline{u}(x,t) < x$. Hence, S is a strict local martingale. ■

3.4.2.2 Illustrative Examples

To illustrate this procedure, we use market price data from the alleged internet dotcom bubble (and beyond), from 1999 to 2005. As explained above, we can use the previous theorem as follows: first we choose a parametric form for the diffusion coefficient and we estimate the parameters as explained in [52] by Genon-Catalot and Jacod after choosing a contrast function to minimize. That is, we choose a parametric form $\sigma(\nu, x)$ where ν is the multidimensional parameter that need to be estimated and a contrast $f(G, x)$. Their estimator is defined as $\hat{\nu}_n = \arg\min \frac{1}{n}\sum_{i=1}^{n} f(\sigma^2(\nu, S_{t_{i-1}}), S_i^n)$ where $S_n^i = \sqrt{n}(S_{t_i} - S_{t_{i-1}})$. Usual choices for the contrast function f in our one dimensional setting are $f_1(G,x) = \ln(G) + \frac{x^2}{G}$ or $f_2(G,x) = (x^2 - G)^2$. We do not provide further details and refer the interested reader to [52] for a detailed description of the estimation procedure.

Then we estimate the volatility function using our non-parametric estimators. If the two volatility function estimates are similar, then by applying the criteria of Theorem 49 to the parametric estimator, we can test for the existence of a price bubble using the comparison theorem as in Theorem 54.

A Conclusive Test: When applied to the stock *Lastminute.com*, the methodology is conclusive.

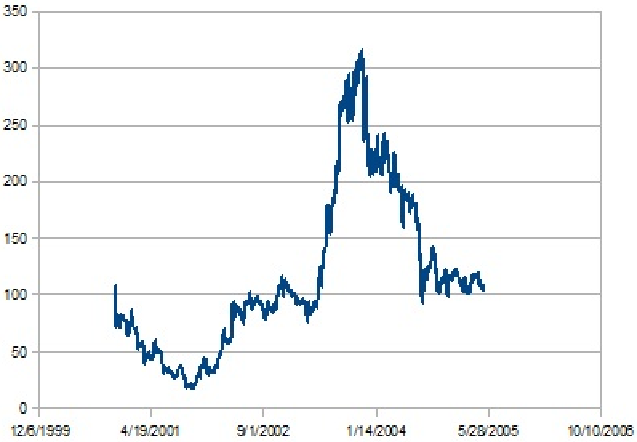

Figure 3.1: Lastminute.com Stock Prices during the alleged Dotcom Bubble.

Intuitively, given the stock price time series as given in Figure 3.1, one suspects the existence of a price bubble. Our test confirms this belief.

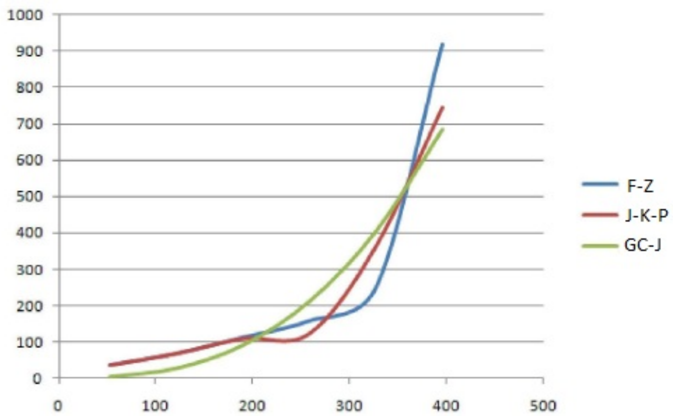

Figure 3.2: Lastminute.com. Estimates of $\sigma(x)$.

This can be seen from Figure 2 above that displays the estimators of Florens-Zmirou (F-Z), the smoothed kernel (J-K-P), and the parametric estimator of Genon-Catalot and Jacod (GC-J), using the power parametric form $\sigma(\sigma_0, \alpha, x) = \sigma_0 x^\alpha$ (here $\nu = (\sigma_0, \alpha)$ is a two-dimensional parameter) and the loglikelihood like contrast $f_1(G, x)$. Using this estimation technique, we find an estimate $\hat{\sigma}(x) = \sigma(\hat{\sigma}_0, \hat{\alpha}, x)$ whose tail behavior leads to the convergence of the integral $\int_\varepsilon^\infty \frac{x}{\hat{\sigma}(x)^2} dx$. Also our estimator (J-K-P) lies above the estimated function (GC-J) hence Theorem 54 guarantess that the price process is a strict local martingale and we have bubble pricing.

An Inconclusive Test: A weakness of this procedure is that the comparison test using parametric estimators might be inconclusive, even when intuitively one suspects a bubble. An example of this phenomenon is that of *Etoys*.

Figure 3.3: eToys Stock Prices during the alleged Dotcom Bubble.

This stock price graph suggests the existence of a price bubble. Our methodology is illustrated in Figure 4.

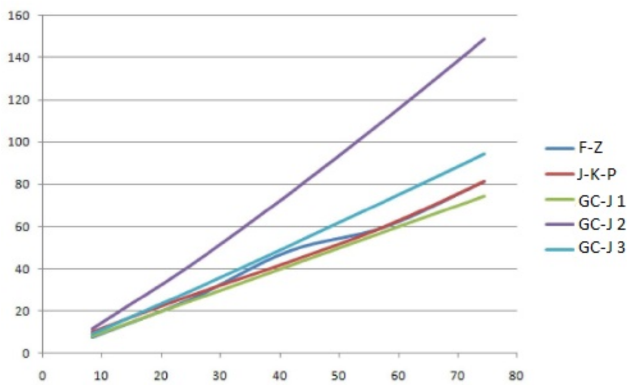

Figure 3.4: eToys. Estimates of $\sigma(x)$.

The three curves included in this figure and labelled GC-J1, GC-J2 and GC-J3 represent the estimators of Genon-Catalot and Jacod, with different parametric forms and contrasts to minimize. GC-J1 and GC-J3 are obtained from the parametric form $\sigma(\sigma_0, \alpha, x) = \sigma_0 x^\alpha$ and the contrasts f_1 and f_2 respectively. GC-J2 is obtained from $\sigma(\sigma_0, \alpha, \beta, x) = \sigma_0 x^\alpha \ln^\beta |x|$ and the contrast f_1. In theory, we have a bubble if $\alpha > 1$ or if $\alpha = 1$ and $\beta > 1$, however the estimated parameters lie in these boundaries values and we see that the curves are so close to linear that we cannot conclude either convergence nor divergence of the integral in Theorem 4. Using this methodology, our test is inconclusive as to whether or not there was a bubble in the stock price of *eToys* during the 1999-2001 period.

3.4.3 Method 2: *RKHS* theory

This section presents our second method for extrapolation. This method is different, in that it is based on *RKHS* theory. Previously in the chapter we have considered parameterized families of functions, so that once the parameter is chosen the tail behavior is determined. We can observe the volatility coefficient σ only on a bounded interval, of course, so it is a leap of faith to assume that (a) it is of the form of the parameterized family of functions considered, and (b) its behavior continues unchanged into the tail. Nevertheless, this is more or less the standard technique in situations such as this.

Our second method is a bit more subtle. Our procedure here consists of two steps:

- We first interpolate an estimate of σ within the bounded interval where we have observations, and in this way we lose the irregularities of non parametric estimators;
- We next extrapolate our function σ by choosing a *RKHS* from a family of Hilbert spaces in such a way as to remain as close as possible (on the bounded interval of observations) to the interpolated function provided in the previous step.

This represents a new methodology which allows us to choose a *good* extrapolation method. We do this via the choice of a certain extrapolating RKHS, which – once chosen – again determines the tail behavior of our volatility σ. If we let $(H_m)_{m\in\mathbb{N}}$ denote our family of RKHS, then any given choice of m, call it m_0, allows us to interpolate *perfectly* the original estimated points, and thus provides a valid *RKHS* H_m with which we extrapolate σ. But this represents a choice of m_0 and not an estimation. So if we stop at this point the method would be as arbitrary as parametric estimation. That is, choosing m_0 is analogous to choosing the parameterized family of functions which fits σ best. The difference is that we do not arbitrarily choose m_0. Instead we choose the index m *given the data available.* In this sense we are using the data twice. To do this we evaluate different RKHSs in order to find the most appropriate one *given the arrangement of the finite number of grid points* from our observations.

The *RKHS* method (see [67]) is intimately related to the reconstruction of functions from scattered data in certain linear functional spaces. The reproducing kernel $Q(x, x')$ that is associated with an *RKHS* $H(\mathcal{D})$ in the spatial domain $\mathcal{D}$, over the coordinate x, is unique and positive and thus constitutes a natural basis for generic interpolation problems.

3.4.3.1 Reproducing Kernel Hilbert Spaces

Let $H(\mathcal{D})$ be a Hilbert space of continuous real valued functions $f(x)$ defined on a spatial domain $\mathcal{D}$. A reproducing kernel Q possesses many useful properties for data interpolation and function approximation problems.

Theorem 55 *There exists a kernel function $Q(x,x^{'})$, the reproducing kernel, in $H(\mathcal{D})$ such that the following properties hold:*

(i) ***Reproducing property.*** *For all x and y,*

$$\begin{aligned} f(x) &= \langle f(x^{'}), Q(x,x^{'})\rangle^{'} \\ Q(x,y) &= \langle Q(x,x^{'}), Q(y,x^{'})\rangle^{'}. \end{aligned}$$

The prime indicates that the inner product $\langle \cdot,\cdot\rangle^{'}$ is performed over $x^{'}$.

(ii) ***Uniqueness.*** *The RKHS $H(\mathcal{D})$ has one and only one reproducing kernel $Q(x,x^{'})$.*

(iii) ***Symmetry and Positivity.*** *The reproducing kernel $Q(x,x^{'})$ is symmetric, i.e. $Q(x^{'},x) = Q(x,x^{'})$, and positive definite, i.e.:*

$$\sum_{i=1}^{n}\sum_{k=1}^{n} c_i Q(x_i,x_k) c_k \geq 0$$

for any set of real numbers c_i and for any countable set of points $(x_i)_{i\in[1,n]}$.

In this framework, interpolation is seen as an inverse problem. The inverse problem is the following. Given a set of real valued data $(f_i)_{i\in[1,M]}$ at M distinct points $S_M = x_i, i \in [1,M]$ in a domain $\mathcal{D}$, and a *RKHS* $H(\mathcal{D})$, find a suitable function $f(x)$ that interpolates these data points. Using the reproducing property, this interpolation problem is reduced to solving the following linear inverse problem :

$$\forall i \in [1,M], f(x_i) = \langle f(x^{'}), Q(x_i,x^{'})\rangle^{'} \tag{3.11}$$

where we need to invert this relation and exhibit the function $f(x)$ in $H(\mathcal{D})$. We refer the reader to [67] for a detailed discussion.

We first present the normal solution that allows an exact interpolation, and second the regularized solution that yields quasi interpolative results, accompanied by an error bound analysis. Then in the next section, we will construct a family of *RKHS*'s that enable us to interpolate not $\sigma(x)$ but $\frac{1}{\sigma(x)^2}$. This transformation makes natural the choice of the family of *RKHS*'s. Note that for every choice of an *RKHS*, on can construct an interpolating function using the input data. For this reason, we define a family of Reproducing Kernel Hilbert Spaces that encapsulate different assumptions on the asymptotic forms and smoothness constraints. From this set, we choose that *RKHS* which best fits the input data in the sense explained below.

Normal Solutions: The most straightforward interpolation approach is to find the normal solution that has the minimal squared norm $||f||^2 = \langle f(x^{'}), f(x^{'})\rangle^{'}$ subject to the interpolation condition (3.11).

That is, given a set of real valued data $\{f_i\}, 1 \leq i \leq K$ specified at K distinct points in a domain $\mathcal{D}$, we wish to find a function f that is the normal solution:

$$f(x) = \sum_{i=1}^{M} c_i Q(x_i, x)$$

where the coefficients c_i satisfy the linear relation :

$$\forall k \in [1, M], \sum_{i=1}^{M} c_i Q(x_i, x_k) = f_k. \tag{3.12}$$

If the matrix Q_M whose entries are the $Q(x_i, x_k)$ is “well conditioned,” then the linear algebraic system above can be efficiently solved numerically. Otherwise, we use regularized solutions.

Regularized Solutions: When the matrix Q_M is "ill conditioned," regularization procedures may be invoked for approximately solving the linear inverse problem. In particular, the Tikhonov regularization procedure produces an approximate solution f_α, which belongs to $H(\mathcal{D})$ and that can be obtained via the minimization of the regularization functional

$$||Qf - F||^2 + \alpha ||f||^2$$

with respect to $f(x)$. Note that here F is the data vector (f_i) and the residual norm $||Qf - F||^2$ is defined as:

$$||Qf - F||^2 = \sum_{i=1}^{M} (\langle f(x^{'}), Q(x_i, x^{'})\rangle^{'} - f_i)^2.$$

The regularization parameter α is chosen to impose a proper balance between the residual constraint $||Qf - F||$ and the magnitude constraint $||f||$. The regularized solution has the form

$$f_\alpha(x) = \sum_{i=1}^{M} c_i^\alpha Q(x_i, x) \tag{3.13}$$

where the coefficients c_i^α satisfy the linear relation:

$$\forall k \in [1, M], \sum_{i=1}^{M} c_i^\alpha (Q(x_i, x_k) + \alpha \delta_{i,k}) = f_k \tag{3.14}$$

where $\delta_{i,k}$ is the Kronecker delta function. Note that for $\alpha > 0$, Q_M^α whose entries are $[Q(x_i, x_k) + \alpha \delta_{i,k}]$ is symmetric and positive definite and the problem can now be solved efficiently. Also, the *RKHS* interpolation method leads to an automatic error estimate of the regularized solution (see [67] for more details).

3.4.3.2 Construction of the Reproducing Kernels

We consider reciprocal power reproducing kernels that asymptotically behave as some reciprocal power of x, over the interval $[0, \infty[$. We are interested in this type of *RKHS* because this is a reasonable assumption for $f(x) = \frac{1}{\sigma^2(x)}$. The CEV model $dS_t = S_t^{\alpha} dW_t$ where $\alpha > 0$ is an important local volatility model proposed in the literature and satisfies this assumption, with $f_{cev}(x) = \frac{1}{x^{2\alpha}}$. We also assume that the function $f(x)$ possesses the asymptotic property

$$\lim_{x \to \infty} x^k f^{(k)}(x) = 0, \forall k \in [1, n-1].$$

for some $n \geq 1$ that controls the minimal required regularity. This property is often satisfied by the volatility functions used in practice. For instance, $x^k f_{cev}^{(k)}(x) = \frac{\prod_{i=0}^{k-1}(-2\alpha - i)}{x^{2\alpha}}$ converges to 0 as x tends to infinity, for all k. This is also satisfied by many volatility functions that explode faster than any power of x, for example $\sigma(x) = x^{\alpha} e^{\beta x}$, with $\alpha > 0$ and $\beta > 0$. The condition appears restrictive only when σ and its derivatives explode too slowly or when σ is bounded, however in these cases, it is likely that there is no bubble and no extrapolation using this *RKHS* theory will be required. We would like to emphasize that the asymptotic property satisfied by f is the key point for the whole method to work as this may be seen from Lemma 91 below.

Concerning the degree of smoothness, we usually take in practice n to be 1, 2 or 3. We can define now our Hilbert space

$$H_n = H_n([0, \infty[) = \left\{ f \in C^n([0, \infty[) \mid \lim_{x \to \infty} x^k f^{(k)}(x) = 0, \forall k \in [1, n-1] \right\}.$$

We now need to define an inner product. A smooth reproducing kernel $q^{RP}(x, x')$ can be constructed via the choice:

$$< f, g >_{n,m} = \int_0^\infty \frac{y^n f^{(n)}(y)}{n!} \frac{y^n g^{(n)}(y)}{n!} \frac{dy}{w(y)}$$

where $w(y) = \frac{1}{y^m}$ is the asymptotic weighting function. From now on we consider the *RKHS* $H_{n,m} = (H_n, <, >_{n,m})$. The next lemma can be shown following the steps in [67].

Lemma 90 *The reproducing kernel is given by*

$$q_{n,m}^{RP}(x, y) = n^2 x_>^{-(m+1)} B(m+1, n) F_{2,1}(-n+1, m+1, n+m+1, \frac{x_<}{x_>})$$

where $x_>$ and $x_<$ are respectively the larger and smaller of x and y, $B(a, b)$ is the beta function and $F_{2,1}(a, b, c, z)$ is Gauss's hypergeometric function.

Proof. From the reproducing property and the explicit use of the inner product defined above, we have

$$f(x) = \langle f(y), q_{n,m}^{RP}(x, y) \rangle = \int_0^\infty \frac{y^n f^{(n)}(y)}{n!} \frac{y^n \partial_{y^n} g_{n,m}^{RP}(y)}{n!} \frac{dy}{w(y)}$$

Then, in the second step, by comparing the equation above with Taylor's formula

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(y)(x-y)^k}{k!} + \frac{1}{(n-1)!} \int_y^x (x-z)^{n-1} f^{(n)}(z) dz$$

with a choice of the reference point $y = \infty$ and the use of the asymptotic property $\lim_{x\to\infty} x^k f^{(k)}(x) = 0$ for all $1 \leq k \leq n-1$, we obtain

$$\frac{y^n}{n!} \frac{\partial^n q_{n,m}^{RP}(x,y)}{\partial y^n} = n\big((y-x)^+\big)^{n-1} y^{-n} w(y)$$

Therefore the corresponding reproducing kernel has the integral representation

$$\begin{aligned} q_{n,m}^{RP}(x,y) &= \langle q_{n,m}^{RP}(x,z), q_{n,m}^{RP}(y,z)\rangle \\ &= n^2 \int_0^\infty \big((z-x)^+\big)^{n-1} \big((z-y)^+\big)^{n-1} z^{-(2n+m)} dz \end{aligned}$$

We need now to carry out the integral above which is equal to

$$\begin{aligned} &\int_{x_>}^\infty (z-x_>)^{n-1}(z-x_<)^{n-1} z^{-(2n+m)} dz \\ &= \int_0^1 (\frac{x_>}{t} - x_<)^{n-1} (\frac{x_>}{t} - x_>)^{n-1} \big(\frac{x_>}{t}\big)^{-(2n+m)} \frac{x_>}{t^2} dt \\ &= \int_0^1 (1 - t\frac{x_<}{x_>})^{n-1} x_>^{2n-2} (1-t)^{n-1} x_>^{-(2n+m)} \frac{x_>}{t^{-m}} dt \\ &= x_>^{-(m+1)} \int_0^1 \frac{(1-t)^{n-1} t^m}{\big(1 - t\frac{x_<}{x_>}\big)^{1-n}} dt \end{aligned}$$

Recall now the closed form of the hypergeometric function

$$F_{2,1}(a,b,c,z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 \frac{t^{b-1}(1-t)^{c-b-1}}{(1-tz)^a} dz$$

By taking $a = -n+1$, $b = m+1$, $c = n+m+1$ and $z = \frac{x_<}{x_>}$, and using the definition of the Beta function, we obtain finally

$$q_{n,m}^{RP}(x,y) = n^2 x_>^{-(m+1)} B(m+1,n) F_{2,1}(-n+1, m+1, n+m+1, \frac{x_<}{x_>})$$

as in the statement of the lemma. ■

Remark 10 *The integers $n-1$ and $m+1$ are respectively the order of smoothness and the asymptotic reciprocal power behavior of the reproducing kernel $q^{RP}(x,y)$. This kernel is a rational polynomial in the variables x and y and has only a finite number of terms, so it is computationally efficient.*

As pointed out above, any choice of n and m creates an *RKHS* $H_{n,m}$ and allows one to construct an interpolating function $f_{n,m}(x)$ with a specific asymptotic behavior. The following result gives the exact asymptotic behavior.

Lemma 91 *For every x, $q^{RP}(x,y)$ is equivalent to $\frac{n^2}{y^{m+1}}B(m+1,n)$ at infinity as a function of y and*

$$\lim_{x\to\infty} x^{m+1} f_\alpha(x) = n^2 B(m+1,n) \sum_{i=1}^{M} c_i^\alpha$$

where f_α is defined as in (3.13) and the constants c_i^α are obtained as in (3.14). Hence, if $\sum_{i=1}^M c_i^\alpha \neq 0$, then $f_\alpha(x)$ is equivalent to $\frac{n^2 B(m+1,n)}{x^{m+1}} \sum_{i=1}^M c_i^\alpha$.

3.4.3.3 Choosing the Best m

The choice of m allows one to decide if the integral in Theorem 49 converges or diverges. If $m > 1$, there is a bubble. This section explains how to choose m. Let us first summarize the idea. We choose the *RKHS* by optimizing over the asymptotic weight m that allows us to construct a function that interpolates the input data points and remains as close as possible to the interpolated function on the finite interval $\mathcal{D}$. This optimization provides an $\overline{m}$ which allows us to construct $\sigma_{\overline{m}}(x)$. We employ a four step procedure:

(i) Non-parametric estimation over $\mathcal{D}$: Estimate $\sigma(x)$ using our non-parametric estimator on a fixed grid $x_1, ..., x_M$ of the bounded interval $\mathcal{D} = [\min S, \max S]$ where $\min S$ and $\max S$ are the minimum and the maximum reached by the stock price over the estimation time interval $[0,T]$. In our illustrative examples, we use the kernel $\phi(x) = \frac{1}{c} e^{\frac{1}{4x^2-1}}$ for $|x| < \frac{1}{2}$, where c is the appropriate normalization constant. The number of data available n and the restriction on the sequence $(h_n)_{n\geq 1}$ makes the number of grid points M relatively small in practice. In our numerical experiments, $7 \leq M \leq 25$.

(ii) Interpolate $\sigma(x)$ over $\mathcal{D}$ using *RKHS* theory: Use any interpolation method on the finite interval $\mathcal{D}$ to interpolate the data points $(\sigma(x_i))_{i\in[1,M]}$. Call the interpolated function $\sigma^b(x)$. For completeness, we provide a methodology to achieve this using the *RKHS* theory. However, any alternative interpolation procedure for a finite interval could be used.

Define the Sobolev space: $H^n(\mathcal{D}) = \left\{u \in L^2(\mathcal{D}) \mid \forall k \in [1,n], u^{(k)} \in L^2(\mathcal{D})\right\}$ where $u^{(k)}$ is the weak derivative of u. The norm that is usually chosen is $||u||^2 = \sum_{k=0}^n \int_{\mathcal{D}} (u^{(k)})^2(x)dx$. Due to Sobolev inequalities, an equivalent and more appropriate norm is $||u|| = \int_{\mathcal{D}} u^2(x)dx + \frac{1}{\tau^{2n}} \int_{\mathcal{D}} (u^{(n)})^2(x)dx$. We denote by $K_{n,\tau}^{a,b}$ the kernel function of $H^n(]a,b[)$, where in this case $\mathcal{D} =]a,b[$. This reproducing kernel is provided for $n=1$ and $n=2$ in the following lemma.

Lemma 92

$$K_{1,\tau}^{a,b}(x,y) = \frac{\tau}{\sinh(\tau(b-a))} \cosh(\tau(b - x_>)) \cosh(\tau(x_< - a))$$
$$K_{2,\tau}^{a,b}(x,y) = L_{x_>}(x_<)$$

and $L_x(t)$ is of the form $\sum_{i=1}^4 \sum_{k=1}^4 l_{ik} b_i(\tau t) b_k(\tau x)$.

We refer to [124, Equation (22) and Corollary 3 on page 28] for explicit analytic expressions for l_{ik} and b_k, which while simple, are nevertheless tedious to write. In both equalities, $x_>$ and $x_<$ respectively stand for the larger and smaller of x and y. In practice, one should check the quality of this interpolation and carefully study the outputs by choosing different τ's before using the interpolated function $\sigma^b = \frac{1}{\sqrt{f^b}}$ in the algorithm detailed above, where $f^b(x) = \sum_{i=1}^M c_i^b K_{n,\tau}^{\mathcal{D}}(x_i, x)$, for all $x \in \mathcal{D}$ and for all $k \in [1, M]$, $\sum_{i=1}^M c_i^b K_{n,\tau}^{\mathcal{D}}(x_i, x_k) = f_k = \frac{1}{\sqrt{\sigma^{est}(x_k)}}$.

(iii) Deciding if an extrapolation is required: If the extended form of the estimated $\sigma(x)$ implies that the volatility does not diverge to ∞ as $x \to \infty$ and remains bounded on $\mathbb{R}^+$, no extrapolation is required. In such a case $\int_\varepsilon^\infty \frac{x}{\sigma^2(x)}$ is infinite and the process is a true martingale. If one decides, however, that $\sigma(x)$ diverges to ∞ as $x \to \infty$, then the next step is required to obtain a 'natural' candidate for its asymptotic behavior as a reciprocal power.

(iv) Extrapolate $\sigma^b(x)$ to $\mathbb{R}^+$ using *RKHS*: Fix $n = 2$ and define

$$\overline{m} = \arg\min_{m \geq 0} \sqrt{\int_{[a,\infty[\cap \mathcal{D}} |\sigma_m - \sigma^b|^2 ds} \tag{3.15}$$

where $f_m = \frac{1}{\sigma_m^2}$ is in the *RKHS* $H_{2,m} = (H_{2,m}([0, \infty[), \langle , \rangle_{RP})$. By definition, all σ_m will interpolate the input data points and $\sigma_{\overline{m}}$ has the asymptotic behavior that best matches our function on the estimation interval. a is the threshold determining closeness to the interpolated function. Choosing a too small is misleading since then it would account more (and unnecessarily) for the interpolation errors over the finite interval $\mathcal{D}$ than desirable. We should choose a large a since we are only interested in the asymptotic behavior of the volatility function. In the illustrative examples below, the threshold a in (3.15) is chosen to be $a = \max S - \frac{1}{3}(\max S - \min S)$.

3.4.4 The dotcom bubble

We illustrate our testing methodology for price bubbles using the stocks that are often alleged ([125] and [110]) as experiencing internet dotcom bubbles. We consider those stocks for which we have tick data. The data was obtained from WRDS [126]. We apply this methodology to four stocks: Lastminute.com, eToys, Infospace, and Geocities. The methodology performs well. The weakness of the method is the possibility of inconclusive tests as illustrated by *eToys*. For *Lastminute.com* and *Infospace* our methodology supports the existence of a price bubble. For *Infospace*, we reproduce the methodology step-by-step. Finally, the study of *Geocities* provides a stock commonly believed to have exhibited a bubble (see for instance [125] and [110]), but for which our method says it did not.

Lastminute.com: Our methodology confirms the existence of a bubble. The stock prices are given in Figure 3.5.

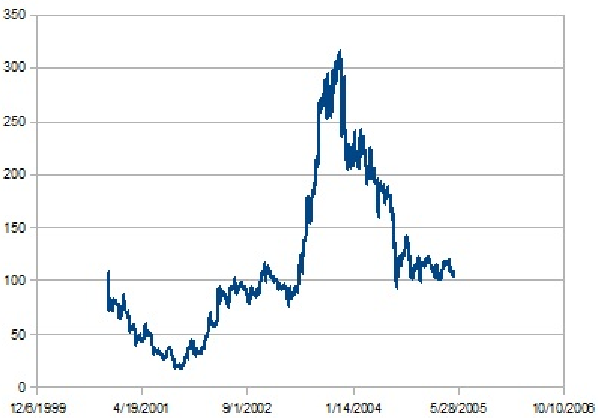

Figure 3.5: Lastminute.com Stock Prices during the alleged Dotcom Bubble.

The optimization performs as expected with the asymptotic behavior given by $\overline{m} = 8.26$, which means that $\sigma(x)$ is equivalent at infinity to a function proportional to x^α with $\alpha = 4.63$. We plot in Figure 3.6 the different extrapolations obtained using different reproducing kernel Hilbert spaces $H_{2,m}$ and their respective reproducing kernels $q^{RP}_{2,m}$.

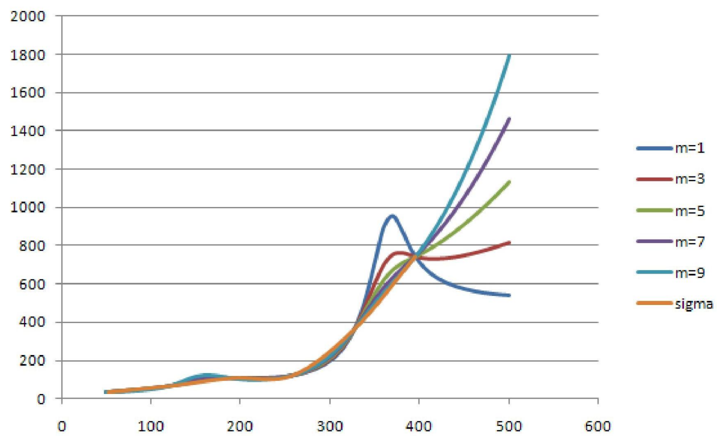

Figure 3.6: Lastminute.com. RKHS estimates of $\sigma(x)$.

Figure 3.6 shows that m is between 7 and 9 as obtained by the optimization procedure. The orange curve labelled (sigma) is the interpolation on the finite interval $\mathcal{D}$ obtained from the non-parametric estimation procedure where the interpolation is achieved using the *RKHS* theory as described in step **(ii)** with the choice of the repoducing kernel Hilbert space $H^1(\mathcal{D})$ and the reproducing kernel $K_{1,6}^{\min S,\max S}$. Then m is optimized as in step **(iv)** so that the interpolating function $\sigma_{\overline{m}}(x)$ is as close as possible to the orange curve in the last third of the domain $\mathcal{D}$, i.e. the threshold a in (3.15) is chosen to be $a =$

$\max S - \frac{1}{3}(\max S - \min S)$.

eToys: While the graph of the stock price of eToys as given in Figure 3.7 makes the existence of a bubble plausible, *the test nevertheless is inconclusive.* Different choices of m giving different asymptotic behaviors are all close to linear (see Figure 3.8).

Figure 3.7: Etoys.com Stock Prices during the alleged Dotcom Bubble.

Because they are so close to being linear, we cannot tell with any level of assurance that the integral in question diverges, or converges. We simply cannot decide which is the case. If it were to diverge we would have a martingale (and hence no bubble), and were it to converge we would have a strict local martingale (and hence bubble pricing).

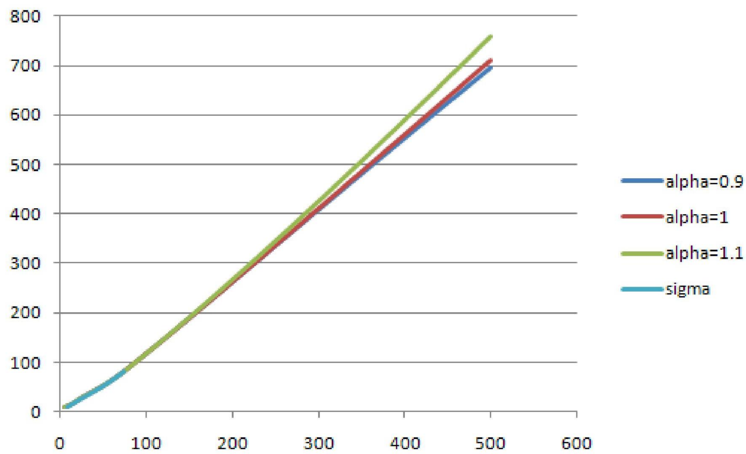

Figure 3.8: eToys. RKHS estimates of $\sigma(x)$.

The estimated $\overline{m}$ is close to one. In Figure 3.8, the powers α are given by $\frac{1}{2}(m+1)$ where m is the weight of the reciprocal power used to define the Hilbert space and its inner product. We plot the extrapolated functions obtained using different Hilbert spaces $H_{2,m}$ together with their reproducing kernels $q^{RP}_{2,m}$. Figure 3.9 shows that the extrapolated functions obtained using these different *RKHS* $H_{2,m}$ produce the same quality of fit on the domain $\mathcal{D}$.

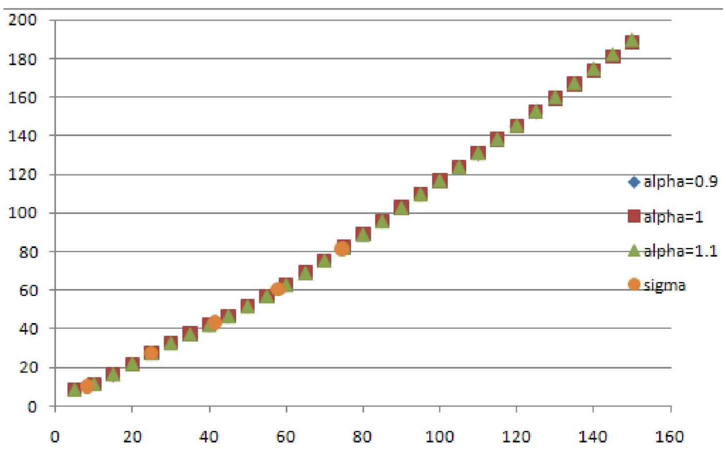

Figure 3.9: eToys. RKHS estimates of $\sigma(x)$, Quality of Fit.

Infospace: Our methodology shows that Infospace exhibited a price bubble. We detail the methodology step by step in this example. The graph of the stock prices in Figure 3.10 suggests the existence of a bubble.

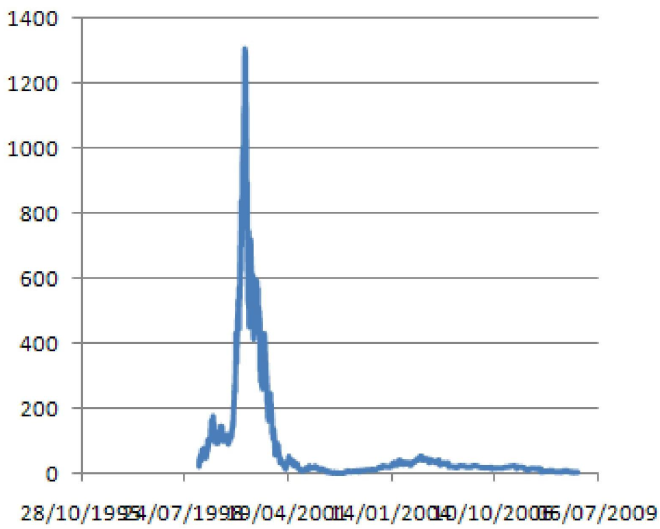

Figure 3.10: Infospace Stock Prices during the alleged Dotcom Bubble.

(i) We compute the Florens-Zmirou's estimator and our smooth kernel local time based estimator, using a sequence $h_n = \frac{1}{n^{\frac{1}{3}}}$. The result is not smooth enough as seen in

Figure 3.11.

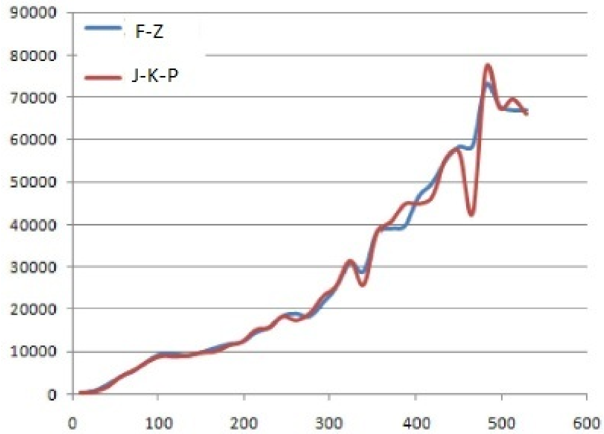

Figure 3.11: Infospace. Non-parametric Estimation using $h_n = \frac{1}{n^{\frac{1}{3}}}$.

(ii) We use the sequence $h_n = \frac{1}{n^{\frac{1}{4}}}$ to compute our estimators (the number of points where the estimation is performed is smaller, M=11). Theoretically, we no longer have the convergence of the Florens-Zmirou's estimator. However, as seen in Figure 3.12, this estimator is robust with respect to the constraint on the sequence h_n. F-Z, LowerBound and UpperBound are Florens-Zmirou's estimator together with the 95% confidence bounds her estimation procedure provides. J-K-P is our estimator.

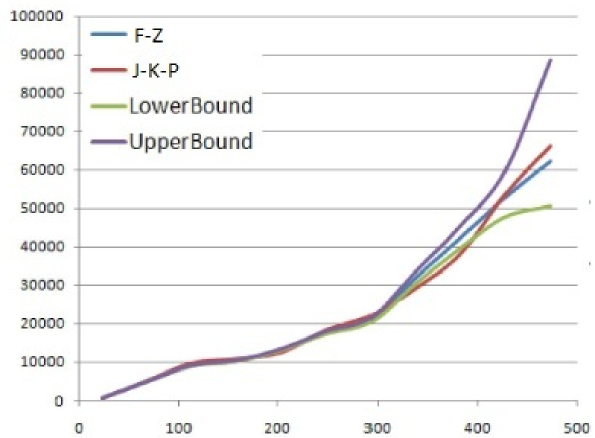

Figure 3.12: Infospace. Non-parametric Estimation using $h_n = \frac{1}{n^{\frac{1}{4}}}$.

(iii) We obtained in (ii) estimations on a fixed grid containing $M = 11$ points, and we now construct a function $\sigma^b(x)$ on the finite domain (see Figure 3.13) which perfectly interpolates those points. Here the *RKHS* used is $H^1(\mathcal{D})$ where $\mathcal{D} = [\min S, \max S]$ together with the reproducing kernels $K^{\mathcal{D}}_{1,\tau}$, where τ takes the values 1, 3, 6 and 9.

The functions obtained using these different reproducing kernels provide the same quality of fit within $\mathcal{D}$ and we can use any of the four outputs as the interpolated function, σ^b, over the finite interval $\mathcal{D}$.

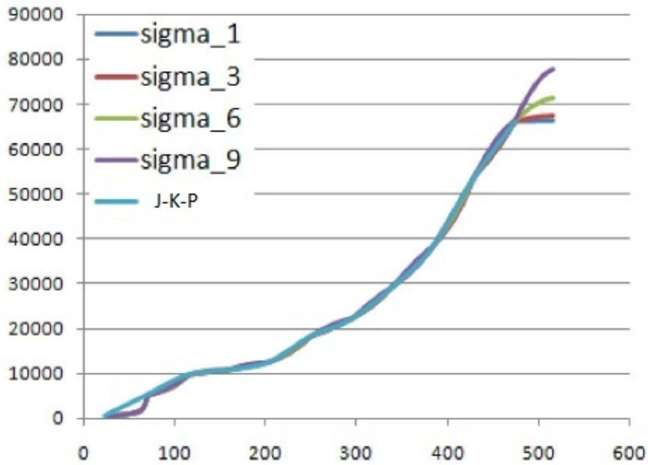

Figure 3.13: Infospace. Interpolation $\sigma^b(x)$ on the compact domain.

(iv) Finally we optimize over m and find the *RKHS* $H_{2,m}$ that allows the best interpolation of the $M = 11$ estimated points and such that the extrapolated function $\overline{\sigma}(x)$ remains as close as possible to $\sigma^b(x)$ on the third right side of $\mathcal{D}$. Of course, the reproducing kernels used in order to construct the functions σ_m and minimize the target error as in (3.15) are $q^{RP}_{2,m}$. We obtain $\overline{m} = 6.17$ (i.e. $\alpha = \frac{\overline{m}+1}{2} = 3.58$) and we can conclude that there is a bubble.

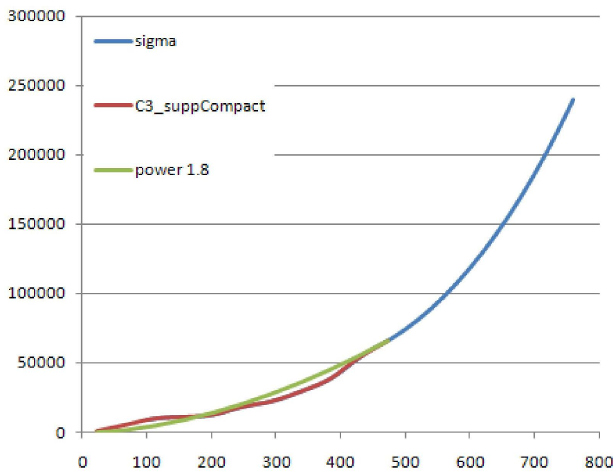

Figure 3.14: Infospace. Final estimator and RKHS Extrapolation.

Remark 11 *One might expect* $\alpha \approx 1.8$ *as suggested by the green curve in Figure 3.14. But this is different from what the RKHS extrapolation has selected. Why? In Figure 3.14, we* *plot the RKHS extrapolation obtained when* $\alpha = 1.8$. *We have proved that*

$$\lim_{x\to\infty} \frac{x^{m+1}}{\overline{\sigma}^2(x)} = 4B(m+1,2)\sum_{i=1}^{M} c_i.$$

The numerical computations give: $\overline{\sigma}(x) \approx \frac{x^{3.58}}{127009}$ *when using optimization over* m *and* $\overline{\sigma}(x) \approx \frac{x^{1.8}}{5.66}$ *when fixing* $\alpha = 1.8$. *Independent of the power chosen, the* c_i*'s and hence the constant of proportionality are automatically adjusted to interpolate the input points. But, as can be seen in Figure 3.15, the power 3.58 is more consistent in terms of extending 'naturally' the behavior of* $\sigma^b(x)$ *to* $\mathbb{R}^+$.

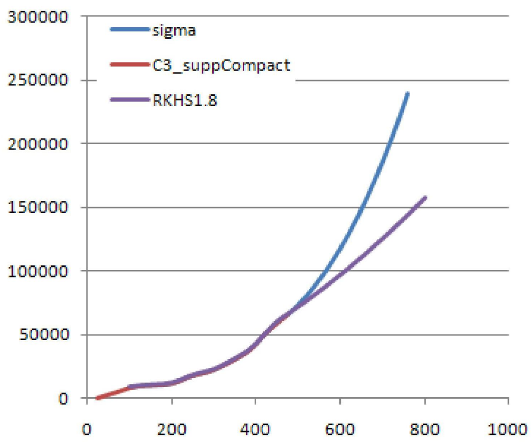

Figure 3.15: Infospace

Geocities: Our methodology shows that this stock did not have a price bubble. The stock prices are graphed in Figure 3.16.

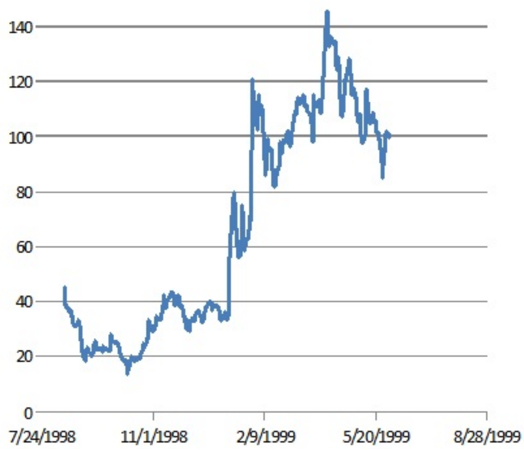

Figure 3.16: Geocities Stock Prices during the alleged Dotcom Bubble.

This is an example where we can stop at step (iii) : we do not need to use *RKHS* theory to extrapolate our estimator in order to determine its asymptotic behavior. As seen from Figure 3.17, the volatility is a nice bounded function, and any natural extension of this behavior implies the divergence of the integral $\int_{\varepsilon}^{\infty} \frac{x}{\sigma^2(x)} dx$. Hence the price process is a true martingale.

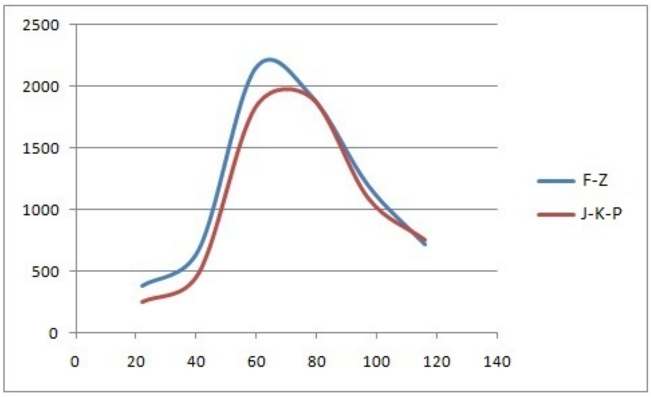

Figure 3.17: Geocities. Estimates of σ.

3.5 Is there a bubble in LinkedIn's stock price? A real case study

Inspired by a New York Times article dated May 23, 2011 [32] discussing whether or not in the aftermath of the LinkedIn IPO the stock price had a bubble, we obtained stock price tick data from Bloomberg (we would like to thank Peter Carr and Arun Verma for help in obtaining quickly the tick data for the stock LinkedIn). And, we used our methodology to test whether LinkedIn's stock price is exhibiting a bubble. *We have found, definitively, that there is a price bubble! The volatility function is well inside the bubble region. There is no doubt about its existence.*

3.5.1 Real time detection : LinkedIn's case

Our results can be summarized in the following graph showing an extrapolation of an estimated volatility function for LinkedIn's stock price.

In Figure 3.25, we have a graph of the volatility coefficient of LinkedIn together with its extrapolation, and because the estimated function increases faster than a straight line, the reader can clearly see that the graph indicates the stock has a price bubble. The blue part of the graph is the estimated function, and the red part is its extrapolation using the technique of Reproducing Kernel Hilbert Spaces (RKHS). These LinkedIn results illustrate the usefulness of our new methodology for detecting bubbles in real time. Our methodology

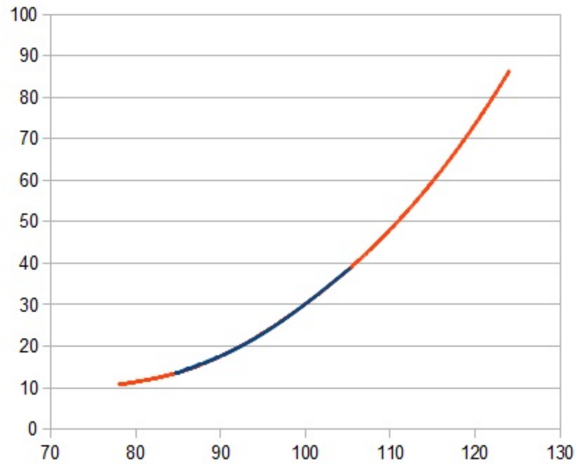

Figure 3.18: Estimation and Extrapolation of the Volatility Function

provides a solution to the problems stated by both Chairman Bernanke and President Dudley, and it is our hope that they will prove useful to regulators, policy makers, and industry participants alike. We apply step by step the methodology derived in the sections above to LinkedIn minute by minute stock price tick data obtained from Bloomberg. We conclude that LinkedIn's stock is indeed experiencing a price bubble.

To perform this validation, we start by assuming that the stock price follows a stochastic differential equation of a form that makes it a diffusion in an incomplete market, namely

$$
\begin{aligned}
dS_t &= \sigma(S_t)dW_t + b(S_t, Y_t)dt && (3.16)\\
dY_t &= s(Y_t)dB_t + g(Y_t)dt && (3.17)
\end{aligned}
$$

where W and B are independent Brownian motions. This model permits that under the physical probability measure, the stock price can have a drift that depends on additional randomness, making the market incomplete. Nevertheless for this family of models, S satisfies the following equation for every neutral measure:

$$dS_t = \sigma(S_t)dW_t$$

Under this evolution, the stock price exhibits a bubble if and only if

$$\int_a^\infty \frac{x}{\sigma^2(x)}dx < \infty \quad for\ all \quad a > 0. \tag{3.18}$$

We test to see if this integral is finite or not. To perform this test, we obtained minute by minute stock price tick data for the 4 business days 5/19/2011 to 5/24/2011 from Bloomberg. There are exactly 1535 price observations in this data set. The time series plot of LinkedIn's stock price is contained in Figure 3.19. The prices used are the open

Figure 3.19: LinkedIn Stock Prices from 5/19/2011 to 5/24/2011.(The observation interval is one minute.)

prices of each minute but the results are not sensitive to using open, high or lowest minute prices instead.

The maximum stock price attained by LinkedIn during this period is \$120.74 and the minimum price was \$81.24. As evidenced in this diagram, LinkedIn experienced a dramatic price rise in its early trading. This suggests an unusually large stock price volatility over this short time period and perhaps a price bubble.

3.5.1.1 Estimating the volatility function

Our bubble testing methodology first requires us to estimate the volatility function σ using local time based non-parametric estimators. We compare the estimation results obtained using both Zmirou's estimator and the estimator developed in Theorem 52. The implementation of these estimators requires a grid step h_n tending to zero. We choose the step size $h_n = \frac{1}{n^{\frac{1}{3}}}$ which implies a grid of 7 points. The statistics for these estimators are displayed in Figure 3.20.

For both tables, the first column contains the stock price grid points. The second column gives the estimated volatility at each grid point. When using Zmirou's estimator, the 95 percent confidence intervals are provided. The graph in Figure 3.20 plots the estimated volatilities for the first 5 grid points together with the confidence intervals.

As computed, the confidence intervals are quite wide. For example, for the first stock price of \$84.665, the volatility estimate using Zmirou's estimator is 19.0354 and the JKP estimator is 13.4404. The Zmirou's 95 percent confidence interval is [\$17.8579, \$20.4816]. This confidence interval does not contain the JKP estimator. In addition, since the neighborhoods of the grid points \$118.915 and \$125.764 are either not visited or visited only once (the last column of the first table), we do not have reliable estimates at these points.

Zmirou's estimator					
x	sigma_Zmirou	lowerBound	upperBound	LocalTime	NbrePoints
84.665	19.0354	17.8579	20.4816	0.0393737	414
91.5149	24.0447	22.6762	25.6951	0.0472675	497
98.3648	22.4606	20.9575	24.3417	0.0330968	348
105.215	37.8995	34.9693	41.7162	0.0239666	252
112.065	86.192	68.0373	137.119	0.00199722	21
118.915	221.362	113.979	1e+006	9.51056e-005	1
125.764	0	0	1e+006	0	0

JKP Estimator		
x	sigma_JKP	LocalTime
84.665	13.4404	0.0619793
91.5149	19.1038	0.0259636
98.3648	27.7474	0.0223718
105.215	38.781	0.0229719
112.065	69.481	0.000708326
118.915	3.95e+014	0
125.764	3.95e+014	0

Figure 3.20: Non-parametric Volatility Estimates.

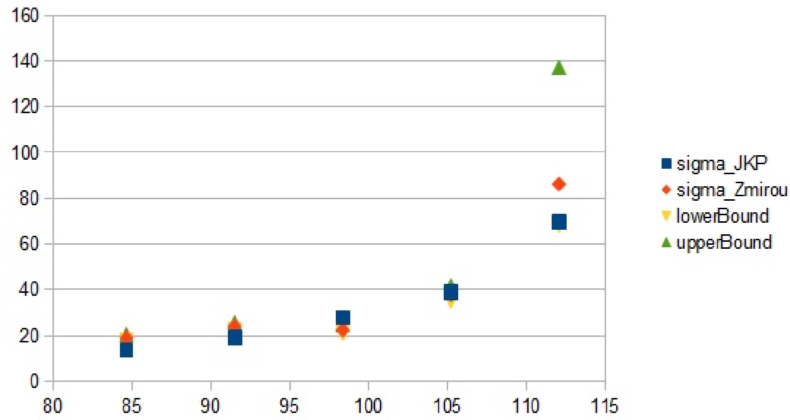

Figure 3.21: Non-parametric Volatility Estimation Results.

Therefore, we restrict ourselves to the grid containing only the first five points. We note that the last point in the new grid $112.065 still has only been visited very few times (21 times to be exact). Given these observations, we apply our methodology twice. In the first test, we use a 5 point grid. In the second test, we remove the fifth point where the estimation is uncertain and we use only a 4 point grid instead.

3.5.1.2 Interpolating the volatility function

The next step in our procedure is to interpolate the shape of the volatility function between these grid points. We use the estimations from our non parametric estimator with the 5 point grid case. For the volatility time scale, we let the 4 day time interval correspond to one unit of time. This scaling does not affect the conclusions we make here. When interpolating one can use any reasonable method. We use both cubic splines and reproducing kernel Hilbert spaces as in item (ii) of our methodology (see 3.4.3.3). The interpolated functions are graphed in Figure 3.22.

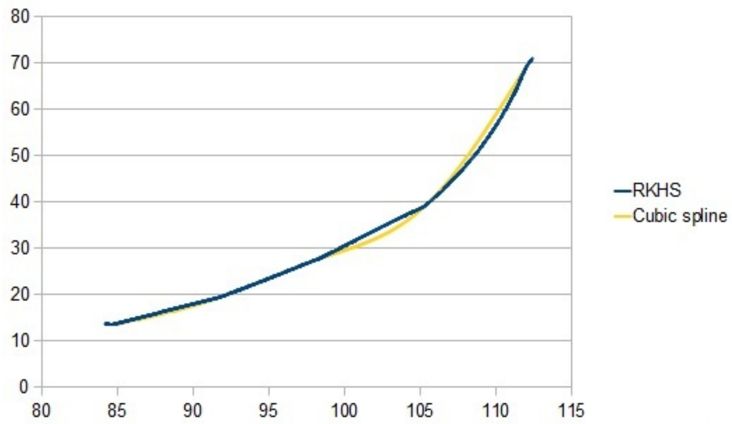

Figure 3.22: Interpolated Volatility using Cubic Splines and the RKHS Theory

Comparing Figures 4 and 5, at least visually, the interpolation appears to provide a reasonable fit for either the cublic spline or the RKHS procedure.

3.5.1.3 Extrapolating the Volatility Function

The next step is to extrapolate the interpolated function σ^b using the RKHS theory to the left and right stock price tails. We construct our extrapolation $\sigma = \sigma_m$ as in item (iv) of 3.4.3.3, by choosing the asymptotic weighting function parameter m such that $f_m = \frac{1}{\sigma_m^2}$ is well-behaved, σ_m exactly matches the points obtained from the non parametric estimation, and σ_m is as close (in norm 2) to σ^b on the last third of the bounded interval where σ^b is defined. We obtain $m = 9.42$. The result is shown in Figure 3.23.

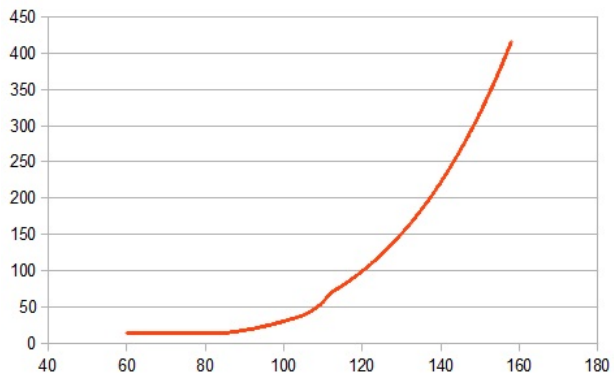

Figure 3.23: RKHS Based Extrapolation of σ^b

Note that the extrapolation contained in Figure 3.23 appears to be a natural extension of that contained in Figure 5. This implies that σ is asymptotically equivalent to a function proportional to x^α with $\alpha = \frac{1+m}{2}$, that is $\alpha = 5.21$. This completes the first approach we used to extrapolate the volatility function's shape to the stock prices outside the fixed grid. We plot below the functions with different asymptotic weighting parameters m obtained using the RKHS extrapolation method, but without optimization. All the functions exactly match the non-parametrically estimated points.

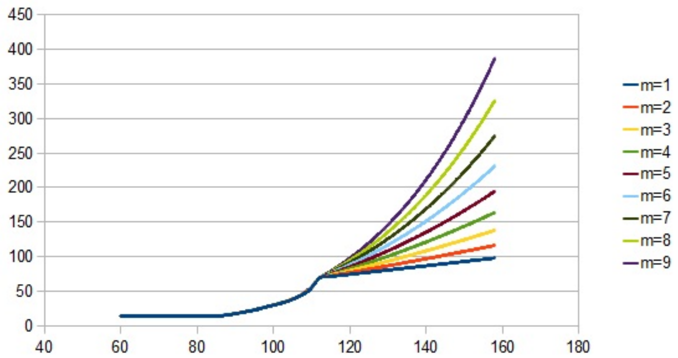

Figure 3.24: Extrapolated Volatility Functions using Different Reproducing Kernels

When we optimized over the weighting functions, we obtained the parameter $m = 9.42$. Visually, this parameter appears in Figure 3.24 to be that estimate most consistent (within all well-behaved volatility functions that exactly match the input data) with a "natural" extension of the behavior of σ^b to $\mathbb{R}^+$. *The power* $\alpha = 5.21$ *implies then that LinkedIn stock price is currently exhibiting a bubble.* Since there is a large standard error for the volatility estimate at the end point $112.065, we remove this point from the grid and repeat our procedure. Also, the rate of increase of the function between the last two last points appears quite large, and we do not want the volatility's extrapolated behavior to follow solely from this fact. Hence, we check to see if we can conclude there is a price bubble based only on the first 4 reliable observation points. We plot in Figure 3.25 the function σ^b (in blue) and its extrapolation to $\mathbb{R}^+$, σ (in red).

With this new grid, we obtain, after optimization, $m = 7.8543$. This leads to the power $\alpha = 4.42715$ for the asymptotic behavior of the volatility. Again, although this power appears to be high, the extrapolated function obtained is the most consistent (within all the well-behaved volatility functions that exactly match the input data) in terms of extending 'naturally' the behavior of σ^b to $\mathbb{R}^+$. Again, we can conclude that there is a stock price bubble. This completes the real time testing and validation of LinkedIn's stock price bubble.

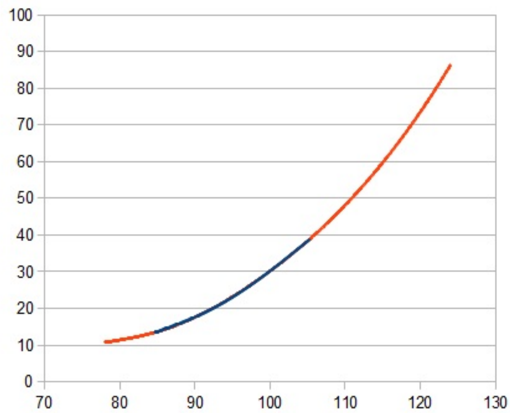

Figure 3.25: RKHS based Extrapolation of σ^b

3.5.2 "Is there a bubble in LinkedIn's stock price?" in the news

The financial press showed a vivid interest to these results during May and June 2011. In this subsection, we attach two articles from financial journals that referenced during the alleged LinkedIn's bubble our work on its real time detection. The hope is that our tools will be of some help to regulators to address those bubbles by recognizing them as they are happening and be perhaps able then to limit their sizes and the damages they usuallly cause when they burst.

3.5.2.1 Reuter's article

Reuter released the article *LinkedIn shares were a bubble: academic model* on June 2, 2011 [9].

"NEW YORK (Reuters) - Three academics say there was a bubble in LinkedIn Corp's shares during the first four days of its trading, which they have determined definitively using a model they designed. The three have devised a model they say can establish in real time whether prices in a market are doomed to collapse. If investors can spot speculative excess in short order, they can avoid overheated markets and better allocate their capital, said Cornell University finance professor Robert Jarrow, who wrote the paper along with Ecole Polytechnique's Younes Kchia and Columbia University's Philip Protter, both mathematicians. "If enough people think there is a bubble and not enough people want to hold it, maybe the bubbles will disappear before they get too large," Jarrow told Reuters. The model, described in a paper currently being peer-reviewed for publication, compares the size of price fluctuations, known as "volatility," with the volatility of a normal stock, which is a stock whose price is what you would pay if you held it forever. If the volatility in the stock you are testing is higher than that of a normal stock, there is a bubble. Take social networking company LinkedIn Corp, for example.The company's shares more than doubled on their first day of trade on the New York Stock Exchange and closed on May 24 – the fourth and final day of trade run through the model – at still more than double their IPO price. The shares have since given up some of their gains, but closed on Wednesday, after just 9 days of trading, at more than 70 percent above their IPO price. Some people would intuitively argue that the speedy share gains indicate a bubble. But the model can prove it: it shows that as LinkedIn's stock price increased, the rate of increase of volatility was abnormally large. The model has been successfully tested against some of the stocks believed to have been bubbles in the 2000 to 2002 dot-com era, according to the paper.

HELP FOR THE FED Investors aren't the only ones who stand to benefit from knowing when bubbles are arising. The Federal Reserve could use its regulatory power to tighten rules for lending in markets that seem to be overheating. Federal Reserve Chairman Ben Bernanke noted in Congressional testimony in 2009 that it is extraordinarily difficult to tell in real time when a bubble is arising, echoing statements from former Fed Chairman Alan Greenspan. The issue even came up, recently, in relation to LinkedIn. When LinkedIn shares jumped 109.4 percent on their first day of trade, Chicago Fed President Charles Evans said he was withholding judgment over whether a new dot-com bubble was under way. "I have no way of knowing that those aren't just exactly the right valuations," Evans told reporters after a speech in Chicago. ”

As a conclusion, if we all agree now that bubbles do exist, and that they can be detected as they happen, the question of interest is what can be done about it? From this perspective, the following article is interesting.

3.5.2.2 Harvard Business Review's article

The Harvard Business Review, released the following article *Can Your Company Survive a Bubble?* on June 9, 2011 [51].

“If you've been wondering whether LinkedIn's stock - selling as I write this for about 72 a share, down from a high of 122.70 during its first day of trading May 19 - is (or at least was) in a bubble, Robert Jarrow, Younes Kchia, and Philip Protter have an answer for you. The trio (respectively, a finance professor at Cornell, an appliedmath Ph.D student at the Ecole Polytechnique, and a statistics professor at Columbia) developed a statistical technique for detecting bubbles that they tested on data from the dot-com heyday. Jarrow, Kchia, and Protter then plugged in LinkedIn's minute-by-minute stock price movements during its first week of trading. The result: "In the case of LinkedIn, the volatility function is well inside the bubble region. There is no doubt about its existence." This is the kind of thing that can drive people outside of quantitative finance a little crazy; there's no reference to company fundamentals, just "sophisticated volatility estimation techniques combined with the method of reproducing kernel Hilbert spaces." Still, it's encouraging to see finance wonks paying serious attention to bubbles, which for decades got almost no attention in academic finance because they weren't supposed to exist. It's also encouraging to see smart people enlisting the methods of other disciplines in the service of bubble detection. Now that we're almost all agreed that bubbles do exist, and that they can to some extent be detected as they happen, the interesting question is what to do about it. There's been lots of debate over how central banks should react to bubbles. But what about at the level of the corporation? What should executives do when their company is caught up in a bubble? As best I can tell, there are three levels of bubble danger (they correlate somewhat with Hyman Minsky's three stages of economic danger: Ponzi, speculative, and hedge finance).

The Ponzi bubble. It's not just an asset-price bubble; it's an asset bubble supported by borrowing that uses those very assets as collateral. And that's not all: Those who've done the borrowing don't have enough income to make the interest payments on their loans; asset prices have to keep rising to keep everything from falling apart. A lot of U.S. mortgage borrowers got into this situation from about 2004 through early 2007. As you may have heard, it ended badly. My sense is that any company that finances itself this way is a fraud, so the best advice for executives is probably turn yourselves in.

The speculative bubble. This is where you keep selling assets (or borrowing against them) to pay your bills, but are working toward a situation where you don't have it anymore. It's okay for you if asset prices stop rising, but if they fall dramatically - or if the market for them simply stops functioning - you're in big trouble. This was to a certain extent what hit Lehman Brothers and Bear Stearns in 2008. They counted on being able to roll over their debt every day, and then suddenly couldn't. But the better example is that of the dot-coms. Dozens of them were able to finance rapid, cash-burning growth in the late 1990s by selling highly priced stock. Then, on March 20, 2000, Barron's published an article listing 51 dot-coms that would run out of cash within 12 months if they weren't able to raise more money or dramatically cut their losses. Market sentiment shifted (possibly because of the Barron's article), the prices of Internet stocks tanked, and within a year or two most of the 51 were no longer with us. This is clearly a risky way to do business. But sometimes it works out: Amazon.com was on that Barron's list of 51 (and its stock price finally climbed back to and surpassed its dot-com era peak in 2009). Markets gave Amazon, an early mover, enough time to build a formidable business, and its management had clearly thought about the possibility of a stock-price drop, and was able to quickly ratchet back spending and turn a profit when it had to. From the evidence of the IPO prospectus it filed last week, Groupon is playing this speculative finance game (it lost 456 million last year, and has 118 million cash in the bank). Does it have a plan for when the flow of cash from investors dries up? The signs aren't all encouraging.

The plain old bubble. LinkedIn is a profitable company. It can stay in business even if its stock price plummets. But dealing with a big stock-price decline, and with the expectations built into an unrealistically high stock price, can be extremely painful. Finance scholar Michael Jensen has argued that overvalued equity was the ruination of Enron and Worldcom, among other companies. But what's a CEO to do about it? I invited Jensen to a Fortune conference a few years back and watched him argue to a roomful of executives that if their companies' stock was was overvalued they should do what they could to bring the price down - by announcing to investors that they thought it was too high, for example. It's fair to say that they all thought Jensen was crazy. A less drastic approach is simply to insulate your company somewhat from financial markets: Don't let stock options account for a very big share of executive pay, focus the company on real performance metrics and not on stock price, build your business around the "real market" of customers rather than the "expectations market" of investors (I stole that terminology from Roger Martin's new Fixing the Game). None of which is going to seem very enticing when your company is caught up in a stock price bubble. "

ACKNOWLEDGMENT.

This chapter is based on a joint work (see [77] and [78]) with Robert Jarrow and Philip Protter.

Chapter 4

Discretely and continuously sampled variance swaps

Contents

4.1 Intoduction

Although variance swaps only started trading in the mid-1900s, they have since become a standard financial instrument useful for managing volatility risk (see Carr and Lee [20] for the history of volatility derivative markets). In the pricing and hedging of variance and volatility swaps, a distinction is made between payoffs that are discretely or continuously sampled. Discretely sampled variance swaps trade in the over-the counter (OTC) markets. In contrast, continuously sampled variance swaps are only an abstract construct, often used to approximate the values of their discretely sampled counterparts (see Broadie and Jain [15],[16], Chan and Platen [24], and Carr and Lee [21]). These approximations are based on the convergence in probability of the discrete payoff as the discrete sampling period goes to zero toward the continuous payoff. But this convergene in probability is weak enough so that the convergence of the expectations might not hold. This approximation can be thought of as an exchange of the expectation operator with the limit in probability operator.

This operator exchange is often invoked without adequate justification. The purpose of this chapter is to characterize the conditions under which this operator exchange is valid.

For our investigation we utilize the martingale pricing methodology where we take as given the asset's price process assuming markets are arbitrage free (in the sense of No Free Lunch with Vanishing Risk (NFLVR)). This evolution is taken to be very general. We only assume that the price process is a strictly positive semimartingale with possibly discontinuous sample paths. We also assume, as a standing hypothesis, that the continuously sampled variance swaps have finite values. Otherwise, before the analysis begins, the approximation would not make sense.

The first two theorems of this chapter (Theorems 56 and 57) characterize the *additional conditions* needed on the price process such that the discretely sampled variance and volatility swaps have finite values. Of course, when the discretely sampled variance swap values do not exist, the approximation is again nonsensical. Surprisingly, we provide examples of otherwise reasonable price processes, with stochastic volatility of the volatility, where these discretely sampled variance swap values do not exist. Next, assuming both the continuous and discretely sampled variance swap values are finite, we study conditions justifying an exchange of the limit and expectation operators. In this regard, under no additional hypotheses, Theorem 60 provides an upper bound for the maximum difference between these two values. Theorem 61 proves, under some additional moment conditions, that the exchange of the two operators is valid. Furthermore, this theorem also provides infomation on the rate of convergence ($1/n$ where n is the number of discretely sampled prices). Lastly, given the recent interest in the $3/2$ stochastic volatility model for pricing volatility derivatives (see Carr and Sun [22], Chan and Platen [24]), we explore its consistency with valuing discretely sampled variance swaps with their continuously sampled counterparts. Here we show that both the discrete and continuously sampled variance swaps have finite values. Unfortunately, we can only prove convergence of the discrete to the continuously sampled variance swap values for some parameter ranges, but not all. A complete characterization of this convergence for the $3/2$ stochastic volatility model remains an open question. An outline for this chapter is as follows. Section 4.2 gives the framework underlying the model. Section 4.3 studies the finiteness and convergence of the discretely sampled variance swap values. Finally, Section 4.4 provides examples to illustrate the theorems proved.

4.2 Framework and mathematical preliminaries

Let a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be given, where the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ satisfies the usual conditions and T is a fixed time horizon. We suppose that there is a liquidly traded asset paying no dividends, whose market price process is modeled by a semimartingale $S = (S_t)_{t\in[0,T]}$ such that $S > 0$ and $S_- > 0$. The value of the money market account is chosen as numeraire, or, viewed differently, the interest rate is zero. The price process is assumed to be arbitrage free in the sense of No Free Lunch with Vanishing Risk (NFLVR), see Delbaen and Schachermayer [33] and [34], which guarantees the existence of at least one equivalent probability measure under which S is a local martingale. We assume P is such a measure, and that the modeler prices future payoffs by taking expectations with respect to P.

4.2.1 Variance swaps

Definition 11 *A variance swap (with strike zero) and maturity T is a contract which pays the "realized variance," i.e. the square of the logarithm returns up to time T, namely*

$$\frac{252}{n}\sum_{i=1}^{n}\left(\ln\frac{S_{t_i}}{S_{t_{i-1}}}\right)^2,$$

where $0 = t_0 < ... < t_n = T$ is a regular sampling of the time interval $[0, T]$, i.e. $t_i - t_{i-1} = \frac{1}{n}$ for $i = 1, \ldots, n$. Finally, 252 is the number of trading days per year. The maturity T is approximately $\frac{n}{252}$.

Rather than considering the quantity

$$P^n(T) = \sum_{i=1}^{n}\left(\ln\frac{S_{t_i}}{S_{t_{i-1}}}\right)^2,$$

practitioners often use its limit

$$P(T) = [\ln S, \ln S]_T$$

in the pricing of variance swaps (see Carr and Lee [20]). That is, one approximates the quantity $E(P^n(T))$ by $E(P(T))$. The continuous approximation $P(T)$ to the variance swap's true payoff $P^n(T)$ is justified by the fact that

$$[\ln S, \ln S]_T = \lim_{n\to\infty}\sum_{i=1}^{n}\left(\ln\frac{S_{t_i}}{S_{t_{i-1}}}\right)^2,$$

where the limit is in probability and taken over a sequence of subdivisions whose mesh size tends to zero.

Analogous to a variance swap, we define the *volatility swap*.

Definition 12 *A volatility swap is a security written on the square root of the variance swap's payoff. We will use the notation*

$$V^n(T) = \sqrt{\sum_{i=0}^{n-1}\left(\ln\frac{S_{t_i}}{S_{t_{i-1}}}\right)^2} \qquad \text{and} \qquad V(T) = \sqrt{\langle \ln S\rangle_T}.$$

Again, $V^n(T)$ is the volatility swap payoff up to multiplicative and additive constants.

Our aim is to investigate the validity of these approximations, motivated by the fact that the convergence of $P^n(T)$ to $P(T)$, respectively $V^n(T)$ to $V(T)$, is only in probability, which *a priori* does not guarantee convergence of their expectations. It is the latter convergence one needs in order to justify the use of this approximation in the context of pricing.

To simplify the notation later on, we introduce the following convention.

Notation. For a process $X = (X_t)_{t\in[0,T]}$ we define

$$\delta_i X = X_{t_i} - X_{t_{i-1}} \qquad (i = 1, \ldots, n).$$

In particular, $P^n(T) = \sum_{i=1}^n (\delta_i \ln S)^2$.

4.2.2 Mathematical preliminaries

Since $S > 0$ and $S_- > 0$ there is a semimartingale M such that

$$S = S_0 \mathcal{E}(M),$$

where $\mathcal{E}(\cdot)$ denotes stochastic exponential; see [76], Theorem II.8.3. In fact, since S is a local martingale and hence a special semimartingale, it follows that from Theorem II.8.21 in [76] that M is a stochastic integral with respect to the local martingale part of S. Furthermore, since $S > 0$, the jumps of M satisfy $\Delta M > -1$, so by the Ansel-Stricker Theorem (see [6]), M is a local martingale.

The following notation is standard. We direct the reader to [76] for details. The random measure μ^M associated with the jumps of M is given by

$$\mu^M(dt, dx) = \sum_s \mathbf{1}_{\{\Delta M_s \neq 0\}} \varepsilon_{(s,\Delta M_s)}(dt, dx),$$

and its predictable compensator is $\nu(dt, dx)$. We may then write

$$S = S_0 \exp\left\{ M - \frac{1}{2}\langle M^c, M^c\rangle - (x - \ln(1+x)) * \mu^M \right\}, \tag{4.1}$$

and furthermore decompose M as

$$M = M^c + M^d = M^c + x * (\mu^M - \nu), \tag{4.2}$$

where M^c is the continuous local martingale part of M and $M^d = x * (\mu^M - \nu)$ is the jump part. Both these processes are local martingales. The following elementary result will be useful later, so we state it as a lemma.

Lemma 93 *The quadratic variation of M and $\ln S$ are given by* *(i)* $[M,M] = \langle M^c, M^c\rangle + x^2 * \mu^M$

(ii) $[\ln S, \ln S] = \langle M^c, M^c\rangle + (\ln(1+x))^2 * \mu^M$

Proof. Part (i) is a standard fact. For part (ii), write $X = (x - \ln(1+x)) * \mu^M$ and notice that $\ln S - \ln S_0 = M - X - \frac{1}{2}\langle M^c, M^c\rangle$. Hence

$$\begin{aligned}[\ln S, \ln S] &= [M - X, M - X] \\ &= [M,M] + [X,X] - 2[M,X] \\ &= \langle M^c, M^c\rangle + x^2 * \mu^M + (x - \ln(1+x))^2 * \mu^M - 2x(x - \ln(1+x)) * \mu^M.\end{aligned}$$

By the Cauchy-Schwartz inequality,

$$\begin{aligned}|x(x - \ln(1+x))| * \mu^M &= \sum_{s\leq \cdot} |\Delta M_s(\Delta M_s - \ln(1 + \Delta M_s))| \\ &\leq \sqrt{\sum_{s\leq\cdot}(\Delta M_s)^2}\sqrt{\sum_{s\leq\cdot}(\Delta M_s - \ln(1+\Delta M_s))^2} < \infty,\end{aligned}$$

so $x(x - \ln(1+x)) * \mu^M$ also converges absolutely, a.s. Termwise manipulation is therefore allowed, so since $x^2 + (x - \ln(1+x))^2 - 2x(x - \ln(1+x)) = (\ln(1+x))^2$, we get

$$x^2 * \mu^M + (x - \ln(1+x))^2 * \mu^M - 2x(x - \ln(1+x)) * \mu^M = (\ln(1+x))^2 * \mu^M.$$

The result follows. ■

We also have the following lemma, which gives the semimartingale decomposition of $\ln S$.

Lemma 94 *Assume that* $\ln S$ *is locally integrable. Then*

$$\ln S - \ln S_0 = M^c + \ln(1+x) * (\mu^M - \nu) - \frac{1}{2}\langle M^c, M^c\rangle - (x - \ln(1+x)) * \nu.$$

Proof. From (4.1) and (4.2) we have

$$\ln S - \ln S_0 = M^c - \frac{1}{2}\langle M^c, M^c\rangle + x * (\mu^M - \nu) - (x - \ln(1+x)) * \mu^M.$$

Our assumption together with the fact that M^c and $x * (\mu^M - \nu)$ are local martingales, hence locally integrable, implies that $\langle M^c, M^c\rangle$ and $(x - \ln(1+x)) * \mu^M$ are locally integrable (notice that both are nonnegative). Hence $(x - \ln(1+x)) * \nu$ is locally integrable (see [76], Proposition II.1.8), so we may add and subtract this quantity to the right side of the previous display to obtain

$$\ln S - \ln S_0 = M^c - \frac{1}{2}\langle M^c, M^c\rangle + \ln(1+x) * (\mu^M - \nu) - (x - \ln(1+x)) * \nu.$$

This is the desired expression. ■

4.3 Approximation using the quadratic variation

4.3.1 Finiteness of expectations

In order for there to be any hope that $E(P(T))$ accurately approximates $E(P^n(T))$, a minimal requirement is that both these quantities be finite. Perhaps somewhat surprisingly, they need not be finite or infinite simultaneously. This is of course a potentially serious issue, since the value of $E(P(T))$, when finite, is nonsensical as an approximation of $E(P^n(T))$ if this is infinite. The following result gives necessary and sufficient conditions for $E(P^n(T))$ to be finite, given that the approximation $P(T)$ is known to have finite expectation.

Theorem 56 *Assume that $P(T) \in L^1$. The following statements are equivalent.*

(i) $P^n(T) \in L^1$ *for at least one* $n \geq 1$

(ii) $\begin{cases} \langle M^c, M^c \rangle_T \in L^2 \\ \\ (x - \ln(1+x)) * \nu_T \in L^2 \end{cases}$

Proof. First of all, note that since $P^n(T) = \sum_{i=1}^n (\delta_i \ln S)^2$, we have $P^n(T) \in L^1$ if and only if $\delta_i \ln S \in L^2$ for each i, which is equivalent to having $\ln S_{t_i} - \ln S_0 \in L^2$ for each i. We thus need to show that this is equivalent to condition (ii) in the statement of the theorem.

By Lemma 93, our basic assumption $P(T) = [\ln S, \ln S]_T \in L^1$ implies that $\langle M^c, M^c \rangle_T \in L^1$ and $(\ln(1+x))^2 * \mu^M_T \in L^1$. Hence both M^c_t and $\ln(1+x) * (\mu^M - \nu)_t$ are in L^2 for every $t \leq T$. Therefore, using the representation from Lemma 94, namely

$$\ln S - \ln S_0 = M^c - \frac{1}{2}\langle M^c, M^c \rangle + \ln(1+x) * (\mu^M - \nu) - (x - \ln(1+x)) * \nu,$$

we deduce that $\ln S_{t_i} - \ln S_0 \in L^2$ for each i holds if and only if

$$\frac{1}{2}\langle M^c, M^c \rangle_{t_i} + (x - \ln(1+x)) * \nu_{t_i} \in L^2$$

for each i. Since both $\langle M^c, M^c \rangle$ and $(x - \ln(1+x)) * \nu$ are nonnegative and nondecreasing, this is equivalent to condition (ii). The proof is finished. ■

We note that the above theorem implies that if $P^n(T) \in L^1$ for some n then $P^n(T) \in L^1$ for all $n \geq 1$. An analogous result holds for volatility swaps.

Theorem 57 *Assume that $V(T) \in L^1$. The following statements are equivalent.*

(i) $V^n(T) \in L^1$ *for some* $n \geq 1$

(ii) $\begin{cases} \langle M^c, M^c \rangle_T \in L^1 \\ \\ (x - \ln(1+x)) * \nu_T \in L^1 \end{cases}$

Proof. Using for instance the fact that all norms on $\mathbb{R}^n$ are equivalent, there are constants $0 < c \leq C < \infty$ such that

$$cV^n(T) \leq \sum_{i=1}^{n} \left|\delta_i \ln S\right| \leq CV^n(T) \quad \text{a.s.}$$

Thus $V^n(T) \in L^1$ is equivalent to $\delta_i \ln S \in L^1$ for each i, which is equivalent to $\ln S_{t_i} - \ln S_0 \in L^1$ for each i.

Now the proof is similar to that of Theorem 56. By Lemma 93, $V(T) = \sqrt{[\ln S, \ln S]_T} \in L^1$ implies that $\sqrt{\langle M^c, M^c\rangle_T}$ and $\sqrt{(\ln(1+x))^2 * \mu_T^M}$ are in L^1, which via the Burkholder-Davis-Gundy inequalities implies that M_t^c and $\ln(1+x) * (\mu^M - \nu)_t$ are in L^1 for each $t \leq T$. By Lemma 94, therefore, $\ln S_{t_i} - \ln S_0 \in L^1$ for each i if and only if

$$\frac{1}{2}\langle M^c, M^c\rangle_{t_i} + (x - \ln(1+x)) * \nu_{t_i} \in L^1$$

for each i. This is equivalent to condition (ii) of the theorem. ■

Again, we note that Theorem 57 implies that if $V^n(T) \in L^1$ for some n then $V^n(T) \in L^1$ for all $n \geq 1$.

When S is continuous we have the following variation of Theorem 56, which indicates that in many cases, $E(P^n(T)) < \infty$ is a stronger requirement than $E(P(T)) < \infty$. This will be used in Section 4.4, where we discuss specific examples. Notice that when S is continuous, $P(T) = [\ln S, \ln S] = \langle M^c, M^c\rangle = \langle M, M\rangle$.

Theorem 58 *Assume that S is continuous. Then the following are equivalent.*

(i) $M_t \in L^1$ *and* $P^n(T) \in L^1$

(ii) $P(T) \in L^2$

Proof. In the continuous case, $\ln S - \ln S_0 = M - \frac{1}{2}\langle M, M\rangle$. First assume (ii), i.e. that $P(T) = \langle M, M\rangle_T \in L^2$. Then certainly $M_t \in L^2$ for every $t \leq T$, hence $\ln S_t \in L^2$ for every $t \leq T$. Therefore $\delta_i \ln S \in L^2$ for every i, and as in the beginning of the proof of Theorem 56, this implies that $P^n(T) \in L^1$. Hence (i) holds. Conversely, if (i) is satisfied, then $\delta_i \ln S \in L^2$ for every i, so $\ln S_T \in L^2$. We also have $M_T \in L^1$ by assumption, so $\langle M, M\rangle_T = 2(M_T - \ln S_T + \ln S_0)$ is in L^1, implying that M_T is in fact in L^2. This lets us strengthen the previous conclusion to $\langle M, M\rangle_T \in L^2$, which is (ii). ■

Notice that the implication $(ii) \Rightarrow (i)$ also follows from Theorem 56. We can prove an analogous result for volatility swaps.

Theorem 59 *Assume that S is continuous. Then the following are equivalent.*

(i) $M_t \in L^1$ *and* $V^n(T) \in L^1$

(ii) $V(T) \in L^2$, *i.e.* $\langle M, M\rangle_T \in L^1$.

Proof. As in the proof of Theorem 57, $V^n(T) \in L^1$ is equivalent to $\delta_i \ln S \in L^1$ for each i. The rest of the proof is similar to that of Theorem 58. ■

In the case of continuous price processes, Theorem 58 makes it clear how one can proceed to construct examples where the approximation $P(T)$ has finite expectation, but the true payoff $P^n(T)$ does not. Indeed, any process S of the form $S = S_0\mathcal{E}(M)$ will do, where M is a continuous local martingale that satisfies

$$\begin{cases} \langle M, M\rangle_T \in L^1 \\ \\ \langle M, M\rangle_T \notin L^2. \end{cases} \tag{4.3}$$

It is clear that such processes exist; what is less clear is to what extent examples can be found among models that appear in applications. In Section 4.4 we provide examples demonstrating that the condition (4.3) can appear in models that may appear innocuous at first sight.

4.3.2 Bounds on the approximation error

In this section we assume that both $E(P(T))$ and $E(P^n(T))$ are finite, and study conditions under which they are close for large n; as already mentioned, although $P^n(T) \to P(T)$ in probability, the expectations need not converge. We start by showing that under general conditions, the two expectations at least cannot be too far apart. We then impose additional structure on the model and give conditions that guarantee convergence. The focus of this section is on variance swaps.

Theorem 60 *Assume that $P^n(T)$ and $P(T)$ both are in L^1. Then there is a constant $C > 0$, independent of n, such that*

$$|E(P(T)) - E(P^n(T))| \leq C \qquad \text{for all } n.$$

The proof of Theorem 60 is straightforward once the following lemma has been established.

Lemma 95 *Assume that $P^n(T)$ and $P(T)$ both are in L^1 and define*

$$\begin{aligned} N &= \ln(1+x) * (\mu^M - \nu) \\ A &= \frac{1}{2}\langle M^c, M^c\rangle + (x - \ln(1+x)) * \nu. \end{aligned}$$

Then $\langle M^c, M^c\rangle_T$ and $[N, N]_T$ are in L^1. Moreover,

$$\big|E\big(P(T)\big) - E\big(P^n(T)\big)\big| \leq E\Big(\sum_{i=1}^n (\delta_i A)^2\Big) + 2E\Big(\sum_{i=1}^n |\delta_i M^c|\delta_i A\Big) + 2E\Big(\sum_{i=1}^n |\delta_i N|\delta_i A\Big),$$

and we have

$$E\Big(\sum_{i=1}^n |\delta_i M^c|\delta_i A\Big) \le \sqrt{E\big(\langle M^c, M^c\rangle_T\big)}\sqrt{E\big(\sum_{i=1}^n (\delta_i A)^2\big)}$$

and

$$E\Big(\sum_{i=1}^n |\delta_i N|\delta_i A\Big) \le \sqrt{E\big([N,N]_T\big)}\sqrt{E\big(\sum_{i=1}^n (\delta_i A)^2\big)}.$$

Proof. Observe that N is a purely discontinuous local martingale with $[N,N] = (\ln(1+x))^2 * \mu^M$, and A is a nondecreasing process. Since by assumption $P(T) = [\ln S, \ln S]_T \in L^1$, Lemma 93 implies that both M^c and N are L^2 martingales, which is the first assertion. Since also $P^n(T) \in L^1$, Theorem 56 shows that $A_T \in L^2$. Again by Lemma 93,

$$\begin{aligned} E\Big([\ln S, \ln S]_T\Big) &= \sum_{i=1}^n E\Big(\delta_i\langle M^c, M^c\rangle\Big) + \sum_{i=1}^n E\Big(\delta_i[N,N]\Big) \\ &= \sum_{i=1}^n E\Big((\delta_i M^c)^2 + (\delta_i N)^2\Big). \end{aligned}$$

Lemma 94 shows that with N and A as above, we have $\ln S - \ln S_0 = M^c + N - A$. Hence $\delta_i \ln S = \delta_i M^c + \delta_i N - \delta_i A$, and therefore

$$\begin{aligned} E\big(P^n(T)\big) - E\big(P(T)\big) &= \sum_{i=1}^n E\Big((\delta_i A)^2 + 2(\delta_i M^c)(\delta_i N) - 2(\delta_i M^c)(\delta_i A) - 2(\delta_i N)(\delta_i A)\Big) \\ &= \sum_{i=1}^n E\Big((\delta_i A)^2 - 2(\delta_i M^c)(\delta_i A) - 2(\delta_i N)(\delta_i A)\Big), \end{aligned}$$

where the second equality holds because M^c and N are orthogonal L^2 martingales. The triangle inequality and Jensen's inequality yield

$$\begin{aligned} \big|E\big(P(T)\big) - E\big(P^n(T)\big)\big| &\le E\Big(\sum_{i=1}^n (\delta_i A)^2\Big) + 2E\Big(\sum_{i=1}^n |\delta_i M^c|\delta_i A\Big) + 2E\Big(\sum_{i=1}^n |\delta_i N|\delta_i A\Big) \\ &= \text{(I)} + \text{(II)} + \text{(III)}. \end{aligned}$$

This settles the first inequality in the statement of the lemma. For (II), applying the Cauchy-Schwartz inequality twice, first for the sum and then the expectation, yields

$$\text{(II)} \le 2E\left(\sqrt{\sum_{i=1}^n (\delta_i M^c)^2}\sqrt{\sum_{i=1}^n (\delta_i A)^2}\right) \le 2\sqrt{E\Big(\sum_{i=1}^n (\delta_i M^c)^2\Big)}\sqrt{E\Big(\sum_{i=1}^n (\delta_i A)^2\Big)}.$$

It only remains to notice that $E\Big(\sum_{i=1}^n (\delta_i M^c)^2\Big) = \sum_{i=1}^n E\Big(\delta_i\langle M^c, M^c\rangle\Big) = E\big(\langle M^c, M^c\rangle_T\big)$. An analogous calculation yields

$$\text{(III)} \le 2\sqrt{E\big([N,N]_T\big)}\sqrt{E\Big(\sum_{i=1}^n (\delta_i A)^2\Big)},$$

thus proving the lemma. ■

Proof of Theorem 60. By Lemma 95 it suffices to bound $E\big(\sum_{i=1}^n (\delta_i A)^2\big)$. Since $\delta_i A \geq 0$ for each i, we get

$$E\Big(\sum_{i=1}^n (\delta_i A)^2\Big) \leq E\Big(\big\{\sum_{i=1}^n \delta_i A\big\}^2\Big) = E(A_T^2) < \infty.$$

The proof is complete. ■

Notice that constant C can be taken to be

$$C = E(A_T^2) + 2\sqrt{E\big(\langle M^c, M^c\rangle_T\big)}\sqrt{E(A_T^2)} + 2\sqrt{E\big([N,N]_T\big)}\sqrt{E(A_T^2)},$$

where N and A are as in Lemma 95. It is worth pointing out that since C is independent of n, it is in particular valid for $n = 1$. It can therefore not be expected to provide a tight bound for large n. Such results can be obtained under additional structure on the model, which we now introduce.

For the remainder of this section we assume that our probability space supports an m-dimensional Brownian motion $W = (W^1, \dots, W^n)$ and a Poisson random measure $\mu = \mu(dt, dz)$ on $\mathbb{R}_+ \times \mathbb{R}$ with intensity measure $dt \otimes F(dz)$, where $\int (z^2 \wedge 1) F(dz) < \infty$. Moreover, we assume that M^c is a stochastic integral with respect to W, and that M^d is a stochastic integral with respect to $\mu - dt \otimes F(dz)$. That is, we assume that there are predictable processes $a^1, \dots, a^m$ and a predictable function $\psi > -1$ such that

$$M_t^c = \sum_{k=1}^m \int_0^t a_s^k dW_s^k \qquad \text{and} \qquad M_t^d = \psi * (\mu - dt \otimes F(dz))_t.$$

In this case the compensator ν of μ^M satisfies

$$\int_0^T \int_{(-1,\infty)} G(s,x)\nu(ds,dx) = \int_0^T \int_{\mathbb{R}} G(s, \psi(s,z)) F(dz) ds$$

for every nonnegative predictable function G. Moreover, it is a classical result that there exists a one-dimensional Brownian motion B and a nonnegative predictable process σ such that

$$M_t^c = \int_0^t \sigma_s dB_s.$$

Under this structure we can formulate conditions on σ and ψ under which the expectation of $P^n(T)$ converges to the expectation of $P(T)$ as $n \to \infty$.

Theorem 61 *With the notation and assumptions just described above, assume that the following conditions hold:*

$$\begin{cases} E\Big\{\int_0^T \sigma_s^4 ds\Big\} < \infty \\[2ex] E\Big\{\int_0^T \int_{\mathbb{R}} (1 \vee |z|^{-p})\big\{\psi(s,z)^2 + (\ln(1+\psi(s,z)))^2\big\} F(dz) ds\Big\} < \infty, \end{cases}$$

for some $p \geq 0$ *such that* $\int_{\mathbb{R}}(1 \wedge |z|^p)F(dz) < \infty$. *Then*

$$\lim_{n\to\infty} |E(P(T)) - E(P^n(T))| = 0.$$

More specifically, there are constants C *and* D *such that*

$$|E(P(T)) - E(P^n(T))| \leq Cn^{-1} + Dn^{-1/2}.$$

Notice that $p \leq 2$ always works, in the sense that $\int_{\mathbb{R}}(1 \wedge z^2)F(dz) < \infty$. However, depending on $F(dx)$, smaller values of p may also work, imposing less stringent restrictions on ψ. In particular, if the Poisson random measure only has finitely many jumps, so that $\int_{\mathbb{R}} F(dx) < \infty$, we may take $p = 0$, and the condition on ψ reduces to

$$E\Big\{\int_0^T \int_{\mathbb{R}} \{\psi(s,z)^2 + (\ln(1+\psi(s,z)))^2\}F(dz)ds\Big\} < \infty.$$

We remark that Theorem 61 generalizes results in [15], where the authors study the Black-Scholes model, the Heston stochastic volatility model, the Merton jump-diffusion model, and the stochastic volatility with jumps model by Bates and Scott.

The proof of Theorem 61 requires two lemmas.

Lemma 96 *For* $0 \leq s < t \leq T$ *we have, a.s.,*

$$\Big\{\int_s^t \int_{(-1,\infty)} (x - \ln(1+x))\nu(du,dx)\Big\}^2$$
$$\leq (t-s)C\int_0^T \int_{\mathbb{R}} (1 \vee |z|^{-p})\{\psi(s,z)^2 + (\ln(1+\psi(s,z)))^2\}F(dz)ds$$

for any $p \geq 0$ *such that* $\int_{\mathbb{R}}(1 \wedge |z|^p)F(dz) < \infty$, *and a finite constant* C *that does not depend on* s, t *or* p.

Proof. Jensen's inequality yields

$$\begin{aligned}\Big\{\int_s^t \int_{(-1,\infty)} &(x - \ln(1+x))\nu(du,dx)\Big\}^2\\ &= \Big\{\int_s^t \int_{\mathbb{R}} (\psi(s,z) - \ln(1+\psi(s,z)))F(dz)ds\Big\}^2\\ &\leq (t-s)\int_s^t \Big\{\int_{\mathbb{R}} G(s,z)F(dz)\Big\}^2 ds,\end{aligned}$$

where we have defined $G(s,z) = \psi(s,z) - \ln(1+\psi(s,z))$. Splitting up the integral over $\mathbb{R}$ as the sum of the integrals over $\{|z| \leq 1\}$ and $\{|z| > 1\}$, and applying the inequality $(a+b)^2 \leq 2a^2 + 2b^2$, we obtain

$$\Big\{\int_{\mathbb{R}} G(s,z)F(dz)\Big\}^2 \leq 2\Big\{\int_{\{|z|\leq 1\}} G(s,z)F(dz)\Big\}^2 + 2\Big\{\int_{\{|z|>1\}} G(s,z)F(dz)\Big\}^2.$$

Since $C_1 = F(\{|z| > 1\}) < \infty$, Jensen's inequality applied to the second term yields

$$\Big\{ \int_{\{|z|>1\}} G(s,z)F(dz) \Big\}^2 \leq C_1 \int_{\{|z|>1\}} G(s,z)^2 F(dz).$$

For the first term, let $\widetilde{F}(dz) = |z|^p F(dz)$. Then $C_2 = \widetilde{F}(\{|z| \leq 1\}) < \infty$ by assumption, and using once again Jensen's inequality we get

$$\begin{aligned}\Big\{ \int_{\{|z|\leq 1\}} G(s,z)F(dz) \Big\}^2 &= \Big\{ \int_{\{|z|\leq 1\}} |z|^{-p} G(s,z)\widetilde{F}(dz) \Big\}^2 \\ &\leq C_2 \int_{\{|z|\leq 1\}} |z|^{-2p} G(s,z)^2 \widetilde{F}(dz) \\ &= C_2 \int_{\{|z|\leq 1\}} |z|^{-p} G(s,z)^2 F(dz).\end{aligned}$$

Assembling the different terms and noting that

$$G(s,z)^2 \leq 2\Big\{ \psi(s,z)^2 + (\ln(1+\psi(s,z)))^2 \Big\}$$

yields the result with $C = 4(C_1 \vee C_2)$. ■

Lemma 97 *Under the assumptions of Theorem 61, we have* $P(T) \in L^1$, $\langle M^c, M^c\rangle_T \in L^2$, *and* $(x - \ln(1+x)) * \nu_T \in L^2$.

Proof. By Jensen's inequality, $E\{(\int_0^T \sigma_s^2 ds)^2\} \leq TE\{\int_0^T \sigma_s^4 ds\}$, which is finite by hypothesis. So $\langle M^c, M^c\rangle_T \in L^2$. To prove that $P(T) \in L^1$ it therefore suffices to note that

$$E\Big\{(\ln(1+x))^2 * \mu_T^M\Big\} = E\Big\{(\ln(1+\psi))^2 * \mu_T\Big\} = E\Big\{(\ln(1+\psi))^2 * (ds \otimes F(dz))\Big\},$$

which is finite by assumption, since $1 \vee |z|^{-p} \geq 1$. Finally, an application of Lemma 96 with $s = 0$ and $t = T$ gives a bound on $E\{((x - \ln(1+x)) * \nu_T)^2\}$ that is finite by assumption. The lemma is proved. ■

Proof of Theorem 61. By Lemma 97, the hypotheses of Lemma 95 are satisfied, so we have the bound

$$\begin{aligned}\big|E\big(P(T)\big) - E\big(P^n(T)\big)\big| &\leq E\Big(\sum_{i=1}^n (\delta_i A)^2\Big) + 2E\Big(\sum_{i=1}^n |\delta_i M^c|\delta_i A\Big) + 2E\Big(\sum_{i=1}^n |\delta_i N|\delta_i A\Big) \\ &= \text{(I)} + \text{(II)} + \text{(III)},\end{aligned}$$

where $N = \ln(1+x) * (\mu^M - \nu)$ and $A = \frac{1}{2}\langle M^c, M^c\rangle + (x - \ln(1+x)) * \nu$. We first deal with (I). Using the inequality $(x+y)^2 \leq 2x^2 + 2y^2$ we obtain

$$\sum_{i=1}^n E\Big((\delta_i A)^2\Big) \leq \sum_{i=1}^n \Big\{ \frac{1}{2} E\Big((\delta_i \langle M^c, M^c\rangle)^2\Big) + 2E\Big((\delta_i (x - \ln(1+x)) * \nu)^2\Big) \Big\}.$$

Now, Jensen's inequality yields

$$E\Big\{(\delta_i\langle M^c, M^c\rangle)^2\Big\} = E\Big\{\Big(\int_{t_{i-1}}^{t_i} \sigma_s^2 ds\Big)^2\Big\} \leq \frac{T}{n}E\Big\{\int_{t_{i-1}}^{t_i} \sigma_s^4 ds\Big\}.$$

Furthermore, Lemma 96 with $s = t_{i-1}$ and $t = t_i$ yields

$$\begin{aligned} &E\Big\{(\delta_i(x - \ln(1+x)) * \nu)^2\Big\} \\ &\qquad \leq \frac{C_1}{n}\int_{t_{i-1}}^{t_i}\int_{\mathbb{R}}(1 \vee |z|^{-p})\{\psi(s,z)^2 + (\ln(1+\psi(s,z)))^2\}F(dz)ds \end{aligned}$$

for some constant C_1 independent of i. Summing over i shows that

$$\begin{aligned} \text{(I)} \leq &\frac{T}{2n}E\Big\{\int_0^T \sigma_s^4 ds\Big\} \\ &+ \frac{2C_1}{n}E\Big\{\int_0^T\int_{\mathbb{R}}(1 \vee |z|^{-p})\{\psi(s,z)^2 + (\ln(1+\psi(s,z)))^2\}F(dz)ds\Big\}, \end{aligned}$$

which is equal to n^{-1} times a constant C that is finite by assumption.

Concerning the two remaining terms (II) and (III), Lemma 95 gives

$$\text{(II)} \leq 2\sqrt{E\big(\langle M^c, M^c\rangle_T\big)}\sqrt{E\Big(\sum_{i=1}^n(\delta_i A)^2\Big)}$$

and

$$\text{(III)} \leq 2\sqrt{E\big([N,N]_T\big)}\sqrt{E\Big(\sum_{i=1}^n(\delta_i A)^2\Big)},$$

so that $\text{(II)} + \text{(III)} \leq C_2\sqrt{\text{(I)}} \leq C_2\sqrt{C}n^{-1/2}$ for some constant C_2. The claim now follows with C as above and $D = C_2\sqrt{C}$. ■

4.4 Examples

4.4.1 Strict local martingales

As it has been extensively emphasized in the previous chapter, the literature on asset price bubbles centers around the phenomenon that S can be a strict local martingale under the risk neutral measure P, see [31] [83], [84]. Moreover, in [43] this issue has been noted to cause complications for pricing using PDE techniques. On the other hand, alternative criteria of no arbitrage type have been proposed by various authors to guarantee the existence of a *true* martingale measure, for instance [122], [19], [65], see for instance Theorems 48 and 49. It is therefore natural to ask about the relationship between our previous results and the true martingale (or strict local martingale) property of S. We give two examples, in the continuous case, showing that the two are not connected in general. More specifically, the examples show that the martingale property of S has little to do with the integrability of $\langle M, M\rangle_T$.

Our first example uses the criterion of Theorem 49. As in Section 4.3.2, B is standard Brownian motion.

Example 10 (S a strict local martingale and $\langle M, M\rangle_T \in L^2$) *Consider the Constant Elasticity of Variance (CEV) models $dS_t = S_t^\alpha dB_t$. By Theorem 49, S is a strict local martingale if and only if $\alpha > 1$. We would like to choose $\alpha > 1$ such that $M_t = \int_0^t S_s^{\alpha-1} dB_s$ is a martingale with an integrable quadratic variation, i.e. $E\{(\int_0^t S_s^{2(\alpha-1)})^2\} < \infty$. This can be achieved with $\varepsilon \in (0,1)$ and $\alpha = 1 + \frac{\varepsilon}{4} > 1$. Indeed,*

$$E\left\{\Big(\int_0^T S_s^{2(\alpha-1)} ds\Big)^2\right\} \leq TE\left(\int_0^T S_s^{4(\alpha-1)} ds\right) = TE\left(\int_0^T S_s^\varepsilon ds\right) = T\int_0^T E(S_s^\varepsilon) ds.$$

Since $\varepsilon \in (0,1)$, $x \mapsto x^\varepsilon$ is concave. Jensen's inequality thus implies that the right side above is dominated by $T\int_0^T E(S_s)^\varepsilon ds \leq T^2 S_0^\varepsilon < \infty$, where $E(S_s) \leq S_0$ because S is a positive local martingale, hence a supermartingale. This shows that S can be a strict local martingale, even if the quadratic variation $\langle M, M\rangle_T = \langle \ln S, \ln S\rangle_T$ is in L^2.

For completeness, we also give a simple example showing that the reverse situation is also possible: that S can be well-behaved (a bounded martingale), while M is not.

Example 11 (S a bounded martingale and M a strict local martingale) *Let X be the reciprocal of a Bessel(3) process. It is well-known that X is a strict local martingale, see e.g. [25, p. 20-21]. Set $M_t = -X_t$, so that M is a strict local martingale with $M_t \in L^1$ and $M_t < 0$ a.s. for all $t \geq 0$. Now, $P(T) = \langle M, M\rangle_T$ is not in L^1 (otherwise M would be a true martingale), and also $P^n(T)$ fails to be in L^1 by Theorem 58.*

However, since $S_t = \mathcal{E}(M)_t = \exp\{-X_t - \frac{1}{2}\langle X, X\rangle_t\} \leq 1$, it is a bounded local martingale, hence a true martingale. In this example, the "bad" behavior of M is caused by its ability to take on very large negative values. This does not carry over to S, since it is obtained through exponentiation.

4.4.2 Stochastic volatility of volatility

We now proceed to give an example of a class of continuous stock price models that look innocuous, but where the conditions (4.3) at the end of Section 4.3.1 are satisfied for certain parameter values. In those cases, Theorem 58 implies that $E(P(T)) < \infty$ but $E(P^n(T)) = \infty$.

We use stochastic volatility models with stochastic volatility of volatility. Let B, W and Z be three Brownian motions, and let ρ denote the correlation between W and Z, i.e. $d\langle W, Z\rangle_t = \rho dt$. No restrictions will be imposed on the correlation structure of

(B, W, Z), other than through the parameter ρ. We consider the following model for the stock price S, its volatility v and the volatility of volatility w:

$$dS_t = S_t\sqrt{v_t}dB_t \tag{4.4}$$
$$dv_t = v_t\sqrt{w_t}dW_t \tag{4.5}$$
$$dw_t = \kappa(\theta - w_t)dt + \eta\sqrt{w_t}dZ_t, \tag{4.6}$$

where κ, θ, η are positive constants. Maintaining our previous notation, where M is the stochastic logarithm of S, we have that $M_t = \int_0^t \sqrt{v_s}dB_s$, so $\langle M, M\rangle_t = \int_0^t v_s ds$. Recall condition (4.3) from Section 4.3.1:

$$\begin{cases} \langle M, M\rangle_T \in L^1 \\ \langle M, M\rangle_T \notin L^2. \end{cases}$$

Note that v is a nonnegative local martingale and hence a supermartingale. Together with Fubini's theorem, this yields

$$E(\langle M, M\rangle_T) = E\Big\{\int_0^T v_s ds\Big\} = \int_0^T E(v_s)ds \le Tv_0,$$

so $\langle M, M\rangle_T \in L^1$. Now we wish to find conditions such that $\langle M, M\rangle_T = \int_0^T v_s ds \notin L^2$. To this end, define

$$T^* = \sup\{t : E(v_t^2) < \infty\},$$

and let $\chi = 2\rho\eta - \kappa$ and $\Delta = \chi^2 - 2\eta^2$. It is proven in [3] that

$$T^* = \begin{cases} \frac{1}{\sqrt{\Delta}}\ln(\frac{\chi+\sqrt{\Delta}}{\chi-\sqrt{\Delta}}) & \text{if } \Delta \ge 0 \text{ and } \chi > 0, \\ \frac{2}{\sqrt{-\Delta}}\Big(\arctan(\frac{\sqrt{-\Delta}}{\chi}) + \pi\mathbf{1}_{\{\chi<0\}}\Big) & \text{if } \Delta < 0, \\ +\infty & \text{otherwise.} \end{cases}$$

The next step is to establish that v is in fact a martingale.

Lemma 98 *The process v defined above is a true martingale.*

Proof. Using Feller's test of explosion (see e.g. [120]), a straightforward calculation shows that w does not explode under P. Therefore, using the same techniques as in [3] or in Section 3.3, it suffices to establish that the auxiliary process $\hat{w}$, defined as the solution to

$$d\hat{w}_t = \big(\kappa(\theta - \hat{w}_t) + \rho\eta\hat{w}_t\big)dt + \eta\sqrt{\hat{w}_t}dZ_t, \qquad \hat{w}_0 = w_0,$$

is non-explosive. This can again be verified using Feller's criterion. ∎

We may now conclude our construction by choosing the parameters ρ, η and κ such that $T^* < \infty$, and then choose $T > T^*$. In this case, Fubini's theorem implies that

$$E\Big\{\Big(\int_0^T v_s ds\Big)^2\Big\} = E\Big\{\int_0^T \int_0^T v_s v_t ds dt\Big\} = \int_0^T \int_0^T E(v_s v_t) ds dt.$$

Moreover, $E(v_s v_t) = E(v_s E(v_t \mid \mathcal{F}_s)) = E(v_s^2)$, so we get

$$E\Big\{\Big(\int_0^T v_s ds\Big)^2\Big\} = \int_0^T \int_0^T E(v_s^2) ds dt \geq \int_{T^*}^T \int_{T^*}^T E(v_s^2) ds dt = \infty.$$

We conclude that $\langle M, M\rangle_T \notin L^2$, and can summarize our findings as follows.

Example 12 *Suppose in the stock price model with stochastic volatility of volatility described above, the parameters are such that $T > T^*$. Then the preceding discussion shows that*

$$E(P(T)) < \infty \qquad \textit{but} \qquad E(P^n(T)) = \infty.$$

That is, the approximation to the variance swap payoff has finite expectation, whereas the true payoff does not.

It is interesting to note that it is sometimes possible to change to an equivalent measure $Q \sim P$, under which the price process is still a local martingale, and such that both $P(T)$ and $P^n(T)$ become integrable. We would like to thank Kerry Back [8] for posing the question of whether or not this can happen. To carry out the construction, let us continue to consider the stochastic volatility of volatility model described above.

Theorem 62 *Assume that we are in the framework of a doubly stochastic volatility model as described in (4.4), (4.5), and (4.6). Suppose that $\Delta \geq 0$ and $\chi > 0$, so that $T^* < \infty$, and assume also that B is independent of (W, Z). Then there is an equivalent measure Q such that S is a local martingale under Q, and $\tilde{T}^* = \sup\{t : E_Q(v_t^2) < \infty\} = \infty$. As a consequence,*

$$E_Q(P(T)) < \infty \qquad \textit{and} \qquad E_Q(P^n(T)) < \infty,$$

and we have $\lim_{n\to\infty} P^n(T) = P(T)$.

Proof. We can find a Brownian motion W', independent of W and B, such that

$$Z_t = \rho W_t + \sqrt{1-\rho^2} W'_t.$$

Let Q be the measure whose density process $Y_t = E_P\big(\frac{dQ}{dP} \mid \mathcal{F}_t\big)$ is given by $dY_t = -Y_t \gamma \sqrt{w_t} dW'_t$, where $\gamma > 0$ is a constant to be determined. To show that Y is indeed a martingale on $[0, T]$, it suffices to verify, as in Lemma 98, that the auxiliary process

$$d\hat{w}_t = \big(\kappa(\theta - \hat{w}_t) - \gamma\rho\eta\hat{w}_t\big)dt + \eta\sqrt{\hat{w}_t} dZ_t, \qquad \hat{w}_0 = w_0,$$

is non-explosive. This can again be done using Feller's criteria. Next, it follows from Girsanov's theorem that the dynamics of w under Q is given by

$$dw_t = \tilde{\kappa}(\tilde{\theta} - w_t)dt + \eta\sqrt{w_t} d\tilde{Z}_t,$$

where

$$\tilde{\kappa} = \kappa + \gamma\eta\sqrt{1-\rho^2}, \qquad \tilde{\theta} = \frac{\kappa\theta}{\kappa + \gamma\eta\sqrt{1-\rho^2}},$$

and $d\tilde{Z}_t = dZ_t + \gamma\sqrt{1-\rho^2}\sqrt{w_t}dt$ is Brownian motion under Q. Hence, if we define

$$\tilde{\chi} = 2\rho\eta - \tilde{\kappa} \qquad \text{and} \qquad \tilde{\Delta} = \tilde{\chi}^2 - 2\eta^2,$$

we have that $\tilde{T}^* = \infty$ if $\tilde{\chi} \leq 0$ and $\tilde{\Delta} \geq 0$. But

$$\tilde{\chi} = \chi - \gamma\eta\sqrt{1-\rho^2}$$

and

$$\tilde{\Delta} = \Delta + \gamma\eta\sqrt{1-\rho^2}\big(\gamma\eta\sqrt{1-\rho^2} - 2\chi\big),$$

so it suffices to choose $\gamma \geq \frac{2\chi}{\eta\sqrt{1-\rho^2}}$.

The verification of the last assertion is straightforward: $E_Q(P(T)) < \infty$ is proved in the same way as under the measure P. To show that $E_Q(P^n(T)) < \infty$ and $\lim_{n\to\infty} P^n(T) = P(T)$, note that $\int_0^T E_Q(v_t^2)dt < \infty$ due to the continuity and finiteness of $E_Q(v_t^2)$ on the compact interval $[0, T]$. An application of Theorem 61 concludes the proof. ■

4.4.3 Time changed geometric Brownian motion

We now consider models obtained as a time change of a geometric Brownian motion. These types of models have appeared frequently in the literature. Let B be standard Brownian motion, and let $(A_t)_{t\in[0,T]}$ be increasing and adapted. We define

$$S_t = S_0\mathcal{E}(M)_t, \qquad \text{where} \qquad M_t = B_{A_t}.$$

In this case we have $\langle M, M\rangle_t = A_t$, so condition (4.3) in Section 4.3.1 amounts to having $A_T \in L^1$ but $A_T \notin L^2$. One such example is of course

$$A_t = \int_0^t v_s ds,$$

with v as in subsection 4.4.2.

4.4.4 The $3/2$-stochastic volatility model

A model that has received considerable attention both in the theoretical and empirical literature is the 3/2-stochastic volatility process. See for example [22] and the references therein. Let B and W be two correlated Brownian motions. The model prescribes the following dynamics for the stock price and its volatility

$$dS_t = S_t\sqrt{v_t}dB_t$$
$$dv_t = v_t(p + qv_t)dt + \varepsilon v_t^{\frac{3}{2}}dW_t$$

where p, q and ε are constants such that $q < \frac{\varepsilon^2}{2}$ and $\varepsilon > 0$. The reason for the upper bound on q is to avoid explosion of v in finite time. To see this, consider the process $R_t = \frac{1}{v_t}$, the reciprocal of v, which satisfies the SDE

$$dR_t = (\varepsilon^2 - q - pR_t)dt - \varepsilon\sqrt{R_t}dW_t.$$

The process R is a square-root process, and it is well-known that this process avoids zero when $\varepsilon^2 - q > \frac{\varepsilon^2}{q}$, which is exactly the condition $q < \frac{\varepsilon^2}{2}$. Let again M be the stochastic logarithm of S, i.e. $M_t = \int_0^t \sqrt{v_s}dB_s$, so that $\langle M, M\rangle_t = \int_0^t v_s ds$. Carr and Sun [22] provide the Laplace transform of the integrated variance $\int_0^T v_s ds$ in closed form.

Lemma 99 *In the* $3/2$*-model, the Laplace transform of the realized variance* $\int_0^T v_s ds$ *is given by*

$$E(e^{-\lambda\int_0^T v_s ds}) = \frac{\Gamma(\gamma-\alpha)}{\Gamma(\gamma)}\left(\frac{2}{\varepsilon^2 y_0}\right)^{\alpha} M(\alpha, \gamma, \frac{-2}{\varepsilon^2 y_0})$$

where $y_0 = v_0\frac{e^{pT}-1}{p}$, $\alpha = -(\frac{1}{2}-\frac{q}{\varepsilon^2})+\sqrt{(\frac{1}{2}-\frac{q}{\varepsilon^2})^2 + 2\frac{\lambda}{\varepsilon^2}}$, $\gamma = 2(\alpha+1-\frac{q}{\varepsilon^2})$, Γ *is the Gamma function, and* M *is the confluent hypergeometric function*

$$M(\alpha, \gamma, z) = \sum_{n=0}^{\infty}\frac{(\alpha)_n}{(\gamma)_n}\frac{z^n}{n!}$$

with the notation $(x)_n = \prod_{i=0}^{n-1}(x+i)$.

Proof. We refer the reader to Carr and Sun [22]. ∎

Since the Laplace transform of the realized variance exists in a neighborhood of zero, all moments of $\int_0^T v_s ds$ are finite. This implies in particular that $E(\langle M, M\rangle_T) \in L^2$. From Theorem 58, both the true variance swap payoff and its approximation have finite expectation.

Recall now the following result proved by Dufresne [39] on the finiteness of moments of the square-root process.

Lemma 100 *Let* $\overline{v} = \frac{2(\varepsilon^2-q)}{\varepsilon^2}$. *Then*

$$\begin{aligned} &\forall\ p < \overline{v}, \qquad E(R_t^{-p}) < \infty \\ &\forall\ p \geq \overline{v}, \qquad E(R_t^{-p}) = \infty \end{aligned}$$

and for all $p \geq -\overline{v}$,

$$E(R_t^p) = \mu_t^p e^{-\lambda_t}\frac{\Gamma(\overline{v}+p)}{\Gamma(\overline{v})}M(\overline{v}+p, p, \lambda_t)$$

where $\mu_t = \frac{\varepsilon^2}{2}\frac{1-e^{-pt}}{p}$, $\lambda_t = \frac{2pv_0}{\varepsilon^2(e^{-pt}-1)}$, Γ *is the Gamma function, and* M *is the congruent hypergeometric function defined in Lemma 99.*

If $q < 0$, define $\kappa = -q > 0$ and $\theta = \frac{p}{\kappa}$. Then the SDE satisfied by v can be re-written as

$$dv_t = \kappa v_t(\theta - v_t)dt + \varepsilon v_t^{\frac{3}{2}} dW_t$$

So under the condition $q < 0$, the process v is mean-reverting with a rate of mean-reversion proportional to v. Also $\overline{v} > 2$ when $q < 0$, so using Lemma 100, it can be seen that $E(v_t^2) = E(R_t^{-2})$ is finite and integrable on $[0, T]$ as a continous function on this compact time interval. Hence the condition of Theorem 61 is satisfied and the expectation of $P^n(T)$ converges to the expectation of $P(T)$ as $n \to \infty$.

Under the condition $0 \leq q < \frac{\varepsilon^2}{2}$, it follows that $1 < \overline{v} \leq 2$ and Lemma 100 implies that $E(v_t^2) = \infty$. By Fubini's theorem, $E(\int_0^T v_s^2 ds) = \infty$ so that the condition of Theorem 61 fails and the convergence of the $E(P^n(T))$ to $E(P(T))$ is not guaranteed anymore.

We now summarize the above findings.

Example 13 *Suppose the stock price follows the* $3/2$*-stochastic volatility model. The above discussion shows that*

(i) Both the true payoff $P^n(T)$ *and the approximation* $P(T)$ *have finite expectation.*

(ii) If $q < 0$*, i.e. when the squared volatility process is mean reverting,* $P^n(T)$ *converges to* $P(T)$ *as* $n \to \infty$*.*

(iii) If $q \geq 0$*, our sufficient condition fails and we can no longer guarantee that* $P^n(T)$ *converges to* $P(T)$*. It is an open problem to establish whether or not this convergence actually takes place.*

ACKNOWLEDGMENT.

This chapter is based on a joint work (see [79]) with Robert Jarrow, Martin Larsson and Philip Protter.

Bibliography

[1] J. Amendinger, Martingale representation theorems for initially enlarged filtrations, *Stochastic Processes and their Applications* Volume 89, Issue 1, Pages 101 - 116, September 2000. (Cited on page 2.)

[2] J. Amendinger, P. Imkeller and M. Schweizer, Additional logarithmic utility of an insider, *Stochastic Processes and their Applications* Volume 75, Issue 2, Pages 263 - 286, July 1998. (Cited on page 2.)

[3] L. Andersen and V. Piterbarg, Moment explosions in stochastic volatility models, *Finance and Stochastics*, 11:29 - 50, 2007. (Cited on page 197.)

[4] S. Ankirchner, On filtration enlargements and purely discontinuous martingales, *Stochastic Processes and their Applications*, Vol. 118, 9, 1662 - 1678, 2008. (Cited on page 4.)

[5] S. Ankirchner, S. Dereich and P. Imkeller, Enlargement of filtrations and continuous Girsanov-type embeddings, *Lecture Notes in Mathematics*, Springer, Vol. 1899, 389 - 410, 2007. (Cited on page 4.)

[6] J. Ansel and C. Stricker. Couverture des actifs contingents et prix maximum, *Annales de l'IHP*, 30(2):303 - 315, 1994. (Cited on page 186.)

[7] F. Antonelli and A. Kohatsu-Higa, Filtration stability of backward SDE's, *Stochastic Analysis and its Applications*, 18, 11 - 37, 2000. (Cited on pages 4 and 17.)

[8] K. Back, Personal communication, 2010. (Cited on page 198.)

[9] C. Baldwin, LinkedIn shares were a bubble: academic model, *Reuters (New York)*, June 2, 2011. Available at http://www.reuters.com/article/2011/06/02/us-bubblemodel-idUSTRE75108X20110602 (Cited on page 179.)

[10] V. Bally and D. Talay, The law of the Euler scheme for stochastic differential equations - Convergence rate of the distribution function, *Probability Theory and Related Fields*, 104, 43 - 60, 1996, Springer-Verlag. (Cited on page 52.)

[11] M. T. Barlow and P. Protter, On convergence of semimartingales, *Séminaire de Probabilités* XXIV, Springer Lecture Notes in Mathematics, 1426, 188 - 193, 1990. (Cited on pages 4 and 40.)

[12] B. Bernanke, Senate Confirmation Hearing, December 2009. This is quoted in many places; one example is *Dealbook, edited by Andrew Ross Sorkin*, January 6, 2010. (Cited on page 137.)

[13] S. Blei and H. J. Engelbert, On exponential local martingales associated with strong Markov continuous local martingales. (Cited on page 138.)

[14] P. Brémaud and M. Yor, Changes of filtrations and probability measures. *Z. Wahrscheinligheitstheorie verw. Gebiete*, 45:269 - 295, 1978. (Cited on page 48.)

[15] M. Broadie and A. Jain, Pricing and hedging volatility derivatives., *Journal of Derivatives*, 15(3):7 - 24, 2008. (Cited on pages 183 and 193.)

[16] M. Broadie and A. Jain, The effect of jumps and discrete sampling on volatility and variance swaps, *International Journal of Theoretical and Applied Finance*, 11 (8), 761 - 797, 2008. (Cited on page 183.)

[17] G. Callegaro, M. Jeanblanc and B. Zargari, Carthaginian enlargement of filtrations, *Prepint* available at http://www.maths.univ-evry.fr/prepubli/306.pdf, October 2010. (Cited on pages 2, 4 and 39.)

[18] L. Campi and U. Çetin, Insider trading in an equilibrium model with default: a passage from reduced-form to structural modeling, *Finance and Stochastics*, 11: 591 - 602, 2007. (Cited on page 2.)

[19] P. Carr, A. Cherny, and M. Urusov, On the martingale property of time-homogeneous diffusions, *Technical report*, Berlin University of Technology, 2009. Available at http://www.math.tu-berlin.de/urusov/papers/cev.pdf. (Cited on pages 138, 142, 145 and 195.)

[20] P. Carr and R. Lee, Volatility derivatives, *Annual Review of Financial Economics*, 1:319 - 339, 2009. (Cited on pages 183 and 185.)

[21] P. Carr and R. Lee, Variance swaps on time-changed Levy processes, *Working paper*, University of Chicago, 2009. (Cited on page 183.)

[22] P. Carr and J. Sun, A new approach for option pricing under stochastic volatility, *Review of Derivatives Research*, 10:87 - 150, 2007. (Cited on pages 184, 199 and 200.)

[23] U. Çetin, R. Jarrow, P. Protter and Y. Yildirim, Modeling credit risk with partial information, *Annals of Applied Probability*, 14 1167 - 1178, 2004. (Cited on page 78.)

[24] L. Chan and E. Platen, Exact pricing and hedging formulas of long dated variance swaps under a 3/2 volatility model, *Working paper*, University of Technology, Sydney, Australia, 2009. (Cited on pages 183 and 184.)

[25] K. L. Chung and R. J. Williams, An Introduction to Stochastic Integration, Second Edition, *Birkhäuser*, Boston, 1990. (Cited on page 196.)

[26] D. Coculescu, H. Geman and M. Jeanblanc, Valuation of default sensitive claims under imperfect information, *Finance and Stochastics*, 12, 195-218, 2008. (Cited on page 79.)

[27] P. Collin-Dufresne, R. Goldstein and J. Helwege, Is credit event risk priced? modeling contagion via the updating of beliefs. *Preprint*, Carnegie Mellon University, 2003. (Cited on pages 78 and 80.)

[28] F. Coquet, J. Mémin and V. Mackevicius, Some examples and counterexamples of convergence of σ-algebras and filtrations, *Lithuanian Mathematical Journal*, 40, No. 3, 2000. (Cited on pages 4, 17, 18, 19, 20 and 21.)

[29] F. Coquet, J. Mémin and L. Slominski, On weak convergence of filtrations, *Séminaire de Probabilités*, XXXV, Springer, 306 - 328, 2004. (Cited on pages 4, 17, 21 and 52.)

[30] J. Corcuera, P. Imkeller, A. Kohatsu-Higa and D. Nualart, Additional utility of insiders with imperfect dynamical information, *Finance and Stochastics*, 8, 437-450 2004. (Cited on page 3.)

[31] A. Cox and D. Hobson, Local martingales, bubbles and option prices, *Finance and Stochastics*, 9:477 - 492, 2005. (Cited on pages 138, 139, 140 and 195.)

[32] J. Creswell, Analysts Are Wary of LinkedIn's Stock Surge, *New York Times Digest*, p. 4, Monday, May 23, 2011. (Cited on page 173.)

[33] F. Delbaen and W. Schachermayer, A general version of the fundamental theorem of asset pricing, *Mathemathische Annalen*, 300, 463 - 520, 1994. (Cited on pages 141 and 185.)

[34] F. Delbaen and W. Schachermayer, Fundamental theorem of asset pricing for unbounded stochastic processes, *Mathematische Annalen*, 312:215 - 250, 1998. (Cited on pages 138, 141, 147 and 185.)

[35] C. Dellacherie and P. A. Meyer, A propos du travail de Yor sur les grossissements des tribus, *Séminaire de Probabilités XII*, Lecture Notes in Mathematics, Springer-Verlag, 649, 69 - 78, 1978. (Cited on pages 13 and 48.)

[36] C. Dellacherie and P. A. Meyer, Probabilities and Potential B, Theory of Martingales, *North-Holland Publishing Company*, 1982. (Cited on page 39.)

[37] C. Dellacherie, B. Maisonneuve and P. A. Meyer, Probabilités et Potentiel, chapitres XVII - XXIV, Processus de Markov (fin). Compléments de calcul stochastique, *Hermann*, Paris, 1992. (Cited on pages 2 and 13.)

[38] D. Duffie and D. Lando, Term structure of credit spreads with incomplete accounting information, *Econometrica* 69, 633 - 664, 2001. (Cited on page 78.)

[39] D. Dufresne, The integrated square-root process, *Research paper no. 90, University of Melbourne*, 2001. (Cited on page 200.)

[40] N. El Karoui, M. Jeanblanc and Y. Jiao, Modelling successive defaults, *Preprint*, 2009. (Cited on pages 80, 105, 107, 108, 109 and 113.)

[41] N. El Karoui, M. Jeanblanc and Y. Jiao, What happens after a default: the conditional density approach, *Stochastic Processes and Their Applications*, Vol. 120, Number 7, 1011-1032, 2010. (Cited on pages 80, 105 and 109.)

[42] N. El Karoui, M. Jeanblanc, Y. Jiao and B. Zargari, Conditional default probability and density, *Preprint* available from http://people.math.jussieu.fr/ jiao/recherche/density.pdf, 2010. (Cited on page 121.)

[43] E. Ekström and J. Tysk, Bubbles, convexity and the black-scholes equation, *Annals of Applied Probability*, 19:1369 - 1384, 2009. (Cited on pages 157 and 195.)

[44] R. Elliott and M. Jeanblanc, Incomplete markets and informed agents, *Mathematical Methods of Operations Research*, 50, 475 - 492, 1998. (Cited on page 2.)

[45] R. Elliott, M. Jeanblanc and M. Yor, On models of default risk, *Mathematical Finance*, 10, 179 - 195, 2000. (Cited on page 78.)

[46] H. J. Engelbert and W. Schmidt, Strong Markov continuous local martingales and solutions of one-dimensional stochastic differential equations, Parts I, II, and III, *Math. Nachr.*, 143, 1989, 167 - 184; 144, 1989, 241 - 281; 151, 1991, 149 - 197. (Cited on page 155.)

[47] K. T. Fang, S. Kotz, and K. W. Ng. Symmetric multivariate and related distributions, *Chapman and Hall, New York*, 1990. (Cited on page 119.)

[48] D. Filipovic, L. P. Hughston and A. Macrina, Conditional density models for asset pricing, *Preprint* available at arXiv:1010.4384v1, 2010. (Cited on page 121.)

[49] D. Florens-Zmirou, On estimating the diffusion coefficient from discrete observations, *Journal of Applied Probability*, 30, 790-804, 1993. (Cited on pages 57, 139 and 148.)

[50] H. Föllmer and P. Protter, Local martingales and filtration shrinkage, *ESAIM: Probability and Statistics*, Available on CJO 2006 doi:10.1051/ps/2010023, 2011. (Cited on pages 3, 8, 26, 29, 38 and 56.)

[51] J. Fox, Can Your Company Survive a Bubble? *Harvard Business Review*, June 9, 2011. Available at http://www.businessweek.com/managing/content/jun2011/ca20110610_077339.htm (Cited on page 180.)

[52] V. Genon-Catalot and J. Jacod, On the estimation of the diffusion coefficient for multi dimensional diffusion processes, *Annales de l'I.H.P*, section B, tome 29, pp.119 - 151, 1993. (Cited on pages 139, 150, 156, 157 and 158.)

[53] K. Giesecke, Successive correlated defaults in a structural model, *Prepint*, 2004. (Cited on page 80.)

[54] K. Giesecke and L. Goldberg, Forecasting default in the face of uncertainty, *Journal of Derivatives* 12, 11 - 25, 2004. (Cited on page 78.)

[55] K. Giesecke and L. Goldberg, Sequential defaults and incomplete information, *Journal of Risk* 7, 1 - 26, 2004. (Cited on page 78.)

[56] K. Giesecke, Correlated default with incomplete information, *Journal of Banking and Finance*, 28, 1521 - 1545, 2004. (Cited on pages 78 and 107.)

[57] K. Giesecke, Default and information, *Journal of Economic Dynamics and Control*, 30, 2281 - 2303, 2005. (Cited on pages 78, 97, 98 and 102.)

[58] J. Goldstein, Interview with William Dudley, the President of the New York Federal Reserve; in *Planet Money*, April 9, 2010. (Cited on page 137.)

[59] X. Guo, R. A. Jarrow and Y. Zeng, Credit risk models with incomplete information, *Preprint* 2008. (Cited on pages 2 and 78.)

[60] X. Guo and Y. Zeng, Intensity process and compensator: a new filtration expansion approach and the Jeulin-Yor theorem, *The Annals of Applied Probability*, Vol. 18, No. 1, 120 - 142, 2008. (Cited on pages 2, 11, 12, 14 and 15.)

[61] S. Heston, M. Loewenstein, and G. A. Willard, Options and bubbles, *Review of Financial Studies*, 20 (2): 359 - 390, 2007. (Cited on page 138.)

[62] C. Hillairet and Y. Jiao, Credit risk with asymmetric information on the default threshold, To appear in *Stochastics and Stochatics Reports*, 2010. (Cited on pages 79, 97 and 98.)

[63] C. Hillairet and Y. Jiao, Information asymmetry in pricing of credit derivatives, *International Journal of Theoretical and Applied Finance*, Vol. 14, No. 5, 611 - 633, 2011. (Cited on page 79.)

[64] P. Hitczenko, Best constants in martingale version of Rosenthal's inequality, *Annals of Probability*, 18, 1656 - 1668, 1990. (Cited on page 154.)

[65] D. Hobson, Comparison results for stochastic volatility models via coupling, *Finance and Stochastics*, 14:129 - 152, 2010. (Cited on page 195.)

[66] M. Hoffmann, L_p Estimation of the diffusion coefficient, *Bernoulli*, 5, 447 - 481, 1999. (Cited on pages 139, 149, 153 and 154.)

[67] T. Hollebeek, T. S. Ho and H. Rabitz, Constructing multidimensional molecular potential energy surfaces from AB Initio data, *Annu. Rev. Phys. Chem.*, 50, 537 - 570, 1999. (Cited on pages 139, 160, 161, 162 and 163.)

[68] D. N. Hoover, Convergence in distribution and Skorohod convergence for thegeneral theory of processes, *Probab. Theory Related Fields*, 89, 239 - 259, 1991. (Cited on page 17.)

[69] K. Y. Hu, A generalization of Kolmogorov's extension theorem and an application to the construction of stochastic processes with random time domains, *The Annals of Probability*, Vol. 16, No. 1, 222 - 230, 1988. (Cited on page 143.)

[70] P. Imkeller, Malliavin's calculus in insider models: additional utility and free lunches, *Mathematical Finance*, 13, 153 - 169, 2003. (Cited on pages 5, 29, 30 and 31.)

[71] P. Imkeller, M. Pontier, and F. Weisz, Free lunch and arbitrage possibilities in a financial market model with an insider, *Stochastic Processes and Applications*, 92, 103 - 130, 2001. (Cited on pages 2, 3, 5, 29, 30, 31 and 62.)

[72] J. Jacod, Grossissement initial, hypothèse ($\mathcal{H}'$) et théorème de Girsanov, *Lecture Notes in Mathematics*, Volume 1118, Springer-Verlag, 1985. (Cited on pages 2, 9, 46 and 79.)

[73] J. Jacod, Rates of convergence to the local time of a diffusion, *Annales de L'Institut Henri Poincaré, Section B*, 34, 505 - 544, 1998. (Cited on page 155.)

[74] J. Jacod, Non-parametric kernel estimation of the coefficient of a diffusion, *Scandinavian Journal of Statistics*, 27, 83 - 96, 2000. (Cited on page 155.)

[75] J. Jacod and P. Protter, Risk neutral compatibility with option prices, *Finance and Stochastics*, 14 (2), 285 - 315, 2010. (Cited on page 141.)

[76] J. Jacod and A. Shiryaev, *Limit Theorems for Stochastic Processes*, Springer-Verlag, second edition, 2003. (Cited on pages 186 and 187.)

[77] R. Jarrow, Y. Kchia and P. Protter, How to detect an asset bubble. *Preprint, Johnson School Research Paper Series No. 28 - 2010*, February 2011. Available at : http://ssrn.com/abstract=1621728 (Cited on page 182.)

[78] R. Jarrow, Y. Kchia and P. Protter, Is there a bubble in LinkedIn's stock price? *To appear in The Journal of Portfolio Management*, May 2011. Available at : http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1858736 (Cited on page 182.)

[79] R. Jarrow, Y. Kchia, M. Larsson and P. Protter, Discretely sampled variance and volatility swaps versus their continuous approximations, *Preprint*, February 2011. Available at : http://arxiv.org/abs/1103.1374 (Cited on page 201.)

[80] R. Jarrow and P. Protter, Large traders, hidden arbitrage and complete markets, *Journal of Banking and Finance*, Vol. 29, 11, 2803-2820, 2005. (Cited on pages 74 and 75.)

[81] R. Jarrow and P. Protter, Structural versus reduced form models: a new information based perspective, *Journal of Investment Management* 2, 1 - 10, 2004. (Cited on page 78.)

[82] R. Jarrow and P. Protter, The martingale theory of bubbles: implications for the valuation of derivatives and detecting bubbles, *Preprint* Electronic copy available at: http://ssrn.com/abstract=1604175, 2010. (Cited on page 140.)

[83] R. Jarrow, P. Protter, and K. Shimbo, Asset price bubbles in a complete market,*Advances in Mathematical Finance*, In Honor of Dilip B. Madan: 105 - 130, 2006. (Cited on pages 138, 139, 140, 141 and 195.)

[84] R. Jarrow, P. Protter, and K. Shimbo, Asset price bubbles in incomplete markets, *Mathematical Finance*, 20(2):145 - 185, 2010. (Cited on pages 138, 139, 140 and 195.)

[85] M. Jeanblanc and Y. Le Cam, Reduced form modelling for credit risk, *Preprint*, November, 2007. (Cited on pages 2, 82 and 84.)

[86] M. Jeanblanc and Y. Le Cam, Progressive enlargement of filtrations with initial times. *Stochastic Processes and Applications*, 119, 2523-2543, 2009. (Cited on pages 2, 3, 26, 30, 80, 107 and 113.)

[87] M. Jeanblanc and M. Rutkowski, Modelling of default risk: an overview,*Mathematical finance: theory and practice*, 2000. (Cited on page 2.)

[88] M. Jeanblanc and S. Song, An explicit model of default time with given survival probability, Preprint, 2010. (Cited on page 78.)

[89] T. Jeulin, Semimartingales et grossissement d'une filtration, *Lecture Notes in Mathematics*, 833, Springer-Verlag, 1980. (Cited on pages 4, 10, 12, 62, 71 and 73.)

[90] T. Jeulin and M. Yor, Grossissement d'une filtration et semi-martingales: formules explicites, *Séminaire de Probabilités XII*, Lecture notes in Mathematics, 649, 78 - 97, Springer 1978. (Cited on page 11.)

[91] T. Jeulin and M. Yor, Inégalité de Hardy, semimartingales et faux-amis, *Séminaire de Probabilités XIII*, Lecture notes in Mathematics, 721, 332 - 359, Springer-Verlag, 1979. (Cited on page 10.)

[92] T. Jeulin and M. Yor, Grossissement de filtrations: examples et applications, *Lecture Notes in Mathematics*, Volume 1118, Springer, Berlin, 1985. (Cited on pages 2 and 10.)

[93] I. Karatzas and S. Shreve, Brownian Motion and Stochastic Calculus, *Springer-Verlag*, 1991. (Cited on pages 145 and 155.)

[94] Y. Kchia and M. Larsson, Credit contagion and risk management with multiple non-ordered defaults, *Preprint*, 2011. Available at : http://arxiv.org/abs/1104.5272 (Cited on page 135.)

[95] Y. Kchia, M. Larsson and P. Protter, Linking progressive and initial filtration expansions, *Preprint*, 2011. Available at : http://arxiv.org/abs/1104.4139 (Cited on page 76.)

[96] Y. Kchia and P. Protter, On progressive filtration expansion with a process, *Preprint*, 2011. Available at : http://arxiv.org/abs/1105.1662 (Cited on page 76.)

[97] Y. Kchia and P. Protter, Toward dynamic models for insider trading, *Working paper*, June 2011. (Cited on page 76.)

[98] A. Kohatsu-Higa, Models for insider trading with finite utility, *Paris-Princeton Lectures on Mathematical Finance*, 2007. (Cited on pages 5, 60 and 61.)

[99] S. Kotani, On a condition that one dimensional diffusion processes are martingales, *In Memoriam Paul-André Meyer*, Séminaire de Probabilités XXXIX, Lecture Notes in Mathematics, 1874, 149 - 156, 2006. (Cited on pages 138, 142 and 145.)

[100] D. Kramkov, Discussion on *How to Detect an Asset Bubble, by R. Jarrow, Y. Kchia, and P. Protter*, Remarks given after the presentation of P. Protter at the meeting *Contemporary Issues and New Directions in Quantitative Finance,* 9^{th} - 11^{th} *July 2010*; Oxford, England, July 10, 2010. (Cited on page 146.)

[101] S. Kusuoka, A remark on default risk models, *Advances in Mathematical Economics*, 1, 69 - 82, 1999. (Cited on pages 2 and 91.)

[102] L. Li and M. Rutkowski, Random times, enlarged filtrations and semimartingale decompositions, *CREST and Sakigake Symposium*, Tokyo, 14 - 18 December 2010. (Cited on page 2.)

[103] D. Lando and M. S. Nielsen, Correlation in corporate defaults: contagion or conditional independence?, November 2009. (Cited on pages 78 and 80.)

[104] A. Lipton and A. Sepp, Credit value adjustment for credit default swaps via the structural default model, *The Journal of Credit Risk*, 5 (2), 123 - 146, 2009. (Cited on page 78.)

[105] A. Lipton and A. Sepp, Multi-factor structural default models and their applications, Working Paper, Bank of America Merrill Lynch, 2009. (Cited on page 78.)

[106] R. S. Liptser and A. Shiryaev, Statistics of Random Process (Second Edition), Springer-Verlag, New York, 2001. (Cited on pages 56 and 88.)

[107] M. Loewenstein and G. A. Willard, Rational equilibrium asset pricing bubbles in continuous trading models, *Journal of Economic Theory*, 91 (1), 17 - 58, 2000. (Cited on page 138.)

[108] M. Loewenstein and G. A. Willard, Local martingales, arbitrage and viability: free snacks and cheap thrills, *Economic Theory*, 16, 135 - 161, 2000. (Cited on pages 138 and 140.)

[109] R. Mansuy and M. Yor, Random times and (Enlargement of) filtrations in a Brownian setting, *Lectures Notes in Mathematics*, Springer, Vol. 1873, 2006. (Cited on page 2.)

[110] J. Markham, A financial history of modern U.S. corporate scandals: from Enron to reform, M.E. Sharpe, 2005. (Cited on pages 139 and 166.)

[111] J. Mémin, Stability of Doob-Meyer decomposition under extended convergence, *Acta Mathematicae Applicatae Sinica, English Series*, 19, 177 - 190, 2003. (Cited on pages 43 and 44.)

[112] A. Mijatovic and M. Urusov, On the martingale property of certain local martingales, To appear in *Probability Theory and Related Fields*, available at arXiv:0905.3701, 2009. (Cited on pages 138, 142, 144 and 145.)

[113] H. Nakagawa, A filtering model on default risk, *Journal Of Mathematical Sciences*, University Of Tokyo, 8(2001), 107 - 142. (Cited on pages 90, 92, 93 and 94.)

[114] A. Nikeghbali, An essay on the general theory of stochastic processes, *Probability Surveys* Vol. 3, 345 - 412, 2006. (Cited on pages 6 and 10.)

[115] A. Nikeghbali, Enlargements of filtrations and path decompositions at non stopping times, *Probability Theory and Related Fields*, Volume 136, Number 4, 524 - 540, 2007. (Cited on page 75.)

[116] A. Nikeghbali and M. Yor, Doob's maximal identity, multiplicative decompositions and enlargements of filtrations, *Illinois J. Math.*, Volume 50, Number 1 - 4, 791 - 814, 2006. (Cited on pages 2 and 83.)

[117] P. Protter, Stochastic Integration and Differential Equations, *Springer-Verlag*, Heidelberg, second edition, 2005. (Cited on pages 2, 6, 10, 11, 15, 22, 46, 51 and 53.)

[118] P. Protter and K. Shimbo, No arbitrage and general semimartingales, *Markov Processes and Related Topics: A Festschrift for Thomas G. Kurtz*, Vol. 4, 267â283, 2008. (Cited on page 62.)

[119] D. Revuz and M. Yor, Continuous Martingales and Brownian Motion, *Springer-Verlag*, Third Edition, 1999. (Cited on page 150.)

[120] L. Rogers and D. Williams, Diffusions, Markov processes and martingales, Volume 2: Itô Calculus, *Cambridge University Press*, 1994. (Cited on page 197.)

[121] P. Schonbucher, Information-driven default contagion, *Preprint*, Department of Mathematics, ETH Zurich, 2004. (Cited on pages 78, 80 and 81.)

[122] C. A. Sin, Complications with stochastic volatility models, *Advances in Applied Probability*, 30(1):256–268, 1998. (Cited on pages 142, 145, 146 and 195.)

[123] Ch. Stricker, Quasi-martingales, martingales locales, semimartingales et filtration naturelle, *Z. Wahr. Verw. Gebiete*, 39, 55-63, 1977. (Cited on page 39.)

[124] C. Thomas-Agnan, Computing a family of reproducing kernels for statistical applications, *Numerical Algorithms*, 13, 21-32, 1996. (Cited on page 166.)

[125] Wikipedia page, http://en.wikipedia.org/wiki/Dot-com_bubble. (Cited on pages 139 and 166.)

[126] WRDS, Lastminute.com, bid prices, from 14 March 2000 to 30 July 2004
WRDS, eToys, bid prices, from 20 May 1999 to 26 February 2001
WRDS, Infospace, bid prices, from 15 December 1998 to 19 September 2002
WRDS, Geocities, bid prices, from 11 August 1998 to 28 May 1999 (Cited on pages 139 and 166.)

[127] M. Yor, Grossissements d'une filtration et semimartingales: théorèmes généraux, *Séminaire de Probabilités XII*, Lecture Notes in Mathematics 649, 61-69, 1978. (Cited on page 10.)

[128] M. Yor, Inégalités entre processus minces et applications. *C.r. Acad. Sci. Paris, Sér. A*, 286, 799-801, 1978. (Cited on page 38.)

[129] M. Yor, Grossissement de filtrations et absolue continuité de noyaux, *Grossissements de filtrations: exemples et applications*, Springer, Editors T. Jeulin and M. Yor, 1985. (Cited on page 3.)

[130] M. Yor, Some aspects of Brownian motion, Part II: Some recent martingale problems, *Lectures in Mathematics*, ETH Zürich, Birkhäuser, Basel, 1997. (Cited on page 2.)